

1
BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

-----o0o-----

TÌM HIỂU MỘT SỐ KỸ THUẬT KHẢM ẢNH

ĐỒ ÁN TỐT NGHIỆP ĐẠI HỌC HỆ CHÍNH QUY

Ngành: Công nghệ Thông tin

HẢI PHÒNG - 2011

MỤC LỤC

MỤC LỤC	1
DANH MỤC HÌNH VẼ	4
DANH MỤC CHỮ VIẾT TẮT	6
MỞ ĐẦU	7
CHƯƠNG 1: KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ KHẢM ẢNH	8
1.1 Khái quát về xử lý ảnh	8
1.1.1 Xử lý ảnh là gì?	8
1.1.2 Các vấn đề cơ bản trong xử lý ảnh	10
1.1.2.1 Một số khái niệm	10
1.1.2.2 Ứng dụng của xử lý ảnh.....	13
1.2 Khảm ảnh	15
1.2.1 Khái niệm khảm ảnh	15
1.2.2 Các kỹ thuật chính được dùng trong khảm ảnh	25
1.2.3 Ứng dụng của khảm ảnh	25
1.2.3.1 Bản đồ số	25
1.2.3.2 Truyền thông quảng cáo	26
CHƯƠNG 2: KỸ THUẬT KHẢM ẢNH	29
2.1 Kỹ thuật đan đa phân giải (Multiresolution spline)	29
2.1.1 Hàm trọng số tương đương	34
2.1.2 Hình chóp Laplace	36
2.1.3 Các điều kiện đường biên	37
2.1.4 Kỹ thuật đan đa phân giải	37
2.1.4.1 Đan chồng các ảnh.....	37
2.1.4.2 Đan các ảnh với các miền tùy ý	39
2.1.4.3 Đan các ảnh không chồng lên nhau	40
2.2 TRÍCH CHỌN ĐẶC TRƯNG TRONG KHẢM ẢNH.....	42

2.2.1	Kỹ thuật trích chọn đặc trưng	42
2.2.1.1	Đặc trưng màu sắc	43
2.2.1.2	Đặc trưng kết cấu	44
2.2.1.3	Đặc trưng hình dạng	45
2.2.1.4	Đặc trưng cục bộ bất biến	45
2.2.2	Ứng dụng kỹ thuật trích chọn đặc trưng trong khám ảnh.....	51
CHƯƠNG 3: CHƯƠNG TRÌNH THỬ NGHIỆM.....		52
3.1	Bài toán	52
3.2	Phân tích, thiết kế.....	52
3.3	Chương trình khám ảnh	53
KẾT LUẬN		58
TÀI LIỆU THAM KHẢO		59
PHỤ LỤC		60

DANH MỤC HÌNH VẼ

Hình 1.1: Quá trình xử lý ảnh	8
Hình 1.2: Các bước cơ bản trong một hệ thống xử lý ảnh.....	8
Hình 1.3: Quan hệ giữa các điểm ảnh.....	11
Hình 1.4: Lược đồ xám của ảnh.....	12
Hình 1.5: Bức ảnh bị dư tối có lược đồ xám tập trung nhiều bên trái	13
Hình 1.6: Hình ảnh sau khi được chỉnh sửa lược đồ xám đã được trải đều.....	13
Hình 1.7: Bản đồ Việt Nam nhận được từ vệ tinh	14
Hình 1.8: Ảnh hồng ngoại.....	15
Hình 1.9: Ảnh con chim được khám từ rất nhiều ảnh nhỏ.....	16
Hình 1.10: Ảnh Panorama được ghép từ 4 hình ảnh chụp liên tiếp.....	17
Hình 1.11: Ảnh polar panorama.....	18
Hình 1.12: Ảnh khám toàn cảnh	19
Hình 1.13 Ảnh mẫu sau khi được xử lý	20
Hình 1.14: Mẫu ảnh khám.....	21
Hình 1.15: Ảnh khám được tạo ra bằng phương pháp trên	22
Hình 1.16: Ảnh góc	23
Hình 1.17: Làm nổi các màu sắc chủ đạo	23
Hình 1.18: Khám các thành phần chủ đạo	24
Hình 1.19: Ảnh hoàn thiện	24
Hình 1.20: Ảnh chụp từ google maps	26
Hình 1.21: Ảnh quảng cáo của hãng IKEA	27
Hình 1.22: Ảnh bìa tạp chí y học Georgetown	28
Hình 2.1: Hai hình ảnh được nối ghép sao cho đường nối càng mịn càng tốt.....	29
Hình 2.2: Các hàm trung bình trọng số và chiều rộng T của miền chuyển tiếp ..	30
Hình 2.3: Một số thử nghiệm kỹ thuật đan ghép với hình ảnh các ngôi sao	31
Hình 2.4: Mô tả phương pháp lọc 1 chiều.	33
Hình 2.5: Hàm trọng số tương đương	35
Hình 2.6: Ảnh ghép từ hai ảnh vệ tinh của San Francisco.....	38
Hình 2.7: Đan hình ảnh con mắt vào trong lòng bàn tay.	40
Hình 2.8: Ứng dụng của việc đan các ảnh không chồng nhau.....	41
Hình 2.9: Biểu đồ mô phỏng việc tính toán các DoG ảnh từ các ảnh kề mờ.....	47
Hình 2.10: Mỗi điểm ảnh được so sánh với 26 láng giềng của nó.	48
Hình 2.11: Quá trình lựa chọn các điểm hấp dẫn.....	48

Hình 2.12: Biểu diễn các vector đặc trưng.....	49
Hình 3.1: Giao diện chương trình khi khởi động.....	53
Hình 3.2: Giao diện chương trình sau khi chọn ảnh nguồn dùng để khám	54
Hình 3.3: Chọn thư mục chứa tập ảnh mẫu.	55
Hình 3.4: Ảnh khám sử dụng ảnh mẫu kích thước 20x20	56
Hình 3.5: Ảnh khám sử dụng ảnh mẫu kích thước 50x50	56
Hình 3.6: Hình ảnh phong cảnh được khám từ tập ảnh mẫu 50x50	57
Hình 3.7: Hình ảnh phong cảnh được khám từ tập ảnh mẫu 20x20	57

DANH MỤC CHỮ VIẾT TẮT

AD	Analog to Digital
Ppi	Pixel per inch
Dpi	Dot per inch
DSLR	Digital Single-lens reflex camera
S.M.A.R.T	Simultaneous Multi-compare Adaptive Rendering Technology
SIFT	Scale Invariant Feature Transform
DoG	Difference-of-Gaussian
DA	Discriminant Analysis
MDA	Mutiple Discriminant Analysis
BDA	biased Discriminant Analysis
SMMS	Symmetric Maximized Minimal Distance in Subspace
JPG	Joint Photographic Experts

MỞ ĐẦU

Trong thời đại hiện nay, công nghệ thông tin có sự phát triển mạnh mẽ về mọi mặt. Cũng chính nhờ sự phát triển này đã góp phần rất lớn trong việc thúc đẩy sự phát triển của các ngành khác như: giáo dục, y tế, quốc phòng an ninh, giải trí, ... Những năm gần đây, phần cứng máy tính đã dần trở nên mạnh mẽ cả về năng lực xử lý lẫn dung lượng lưu trữ vì thế mà các lĩnh vực phần mềm cũng được phát triển theo đặc biệt là lĩnh vực xử lý ảnh.

Ngày nay, hầu hết các loại máy ảnh thông thường dù độ phân giải có cao nhưng cũng chỉ ghi lại được một phần của những đối tượng lớn ví dụ như sân vận động, công viên, thành phố hay thậm chí là mặt trăng hoặc một hành tinh nào đó. Hoặc tạo ra các hình ảnh độc đáo từ rất nhiều các ảnh nhỏ, điều này thì máy ảnh khó có thể thực hiện được. Chính vì vậy mà chúng ta cần đến kĩ thuật khảm ảnh. Khảm ảnh có thể xét trên hai phương diện đó là khảm ảnh toàn cảnh và khảm ảnh nhiều lớp. Mục đích của khảm ảnh chính là việc tạo ra những bức ảnh có tầm nhìn rộng hơn hoặc những bức hình nghệ thuật độc đáo.

Trên cơ sở đó em đã lựa chọn đề tài: **“Tìm hiểu một số kỹ thuật khảm ảnh”** với mục đích chính là tìm hiểu một số kỹ thuật khảm ảnh đồng thời cài đặt một chương trình thử nghiệm.

Về lý thuyết:

- Tìm hiểu khái quát về xử lý ảnh và một số kỹ thuật khảm ảnh.
- Tìm hiểu một số kỹ thuật khảm ảnh trong xử lý ảnh.

Về thực tiễn:

- Cài đặt thử nghiệm một trong những kỹ thuật tìm hiểu được.

Cấu trúc chính của đồ án bao gồm 3 chương:

Chương 1: Khái quát về xử lý ảnh và khảm ảnh

Trình bày khái quát về xử lý ảnh và khảm ảnh.

Chương 2: Kỹ thuật khảm ảnh

Trình bày một số kỹ thuật khảm ảnh phổ biến.

Chương 3: Chương trình thử nghiệm

Chương trình ứng dụng và một số kết quả thu được.

CHƯƠNG 1: KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ KHẢM ẢNH

1.1 Khái quát về xử lý ảnh

1.1.1 Xử lý ảnh là gì?

Con người thu nhận thông tin qua các giác quan, trong đó thị giác đóng vai trò quan trọng nhất. Những năm trở lại đây với sự phát triển của phần cứng máy tính, xử lý ảnh và đồ hoạ đã phát triển một cách mạnh mẽ và có nhiều ứng dụng trong cuộc sống. Xử lý ảnh và đồ hoạ đóng một vai trò quan trọng trong tương tác người máy.

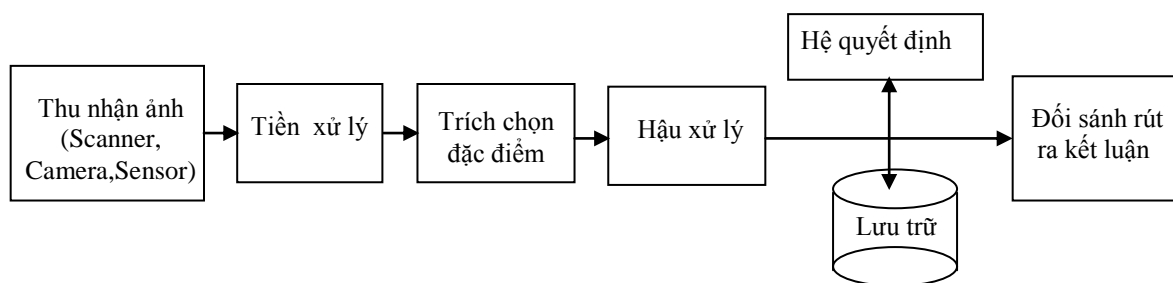
Quá trình xử lý ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho ra kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể là một ảnh “tốt hơn” hoặc một kết luận[1].



Hình 1.1: Quá trình xử lý ảnh

Ảnh có thể xem là tập hợp các điểm ảnh và mỗi điểm ảnh được xem như là đặc trưng cường độ sáng hay một dấu hiệu nào đó tại một vị trí nào đó của đối tượng trong không gian và nó có thể xem như một hàm n biến $P(c_1, c_2, \dots, c_n)$. Do đó, ảnh trong xử lý ảnh có thể xem như ảnh n chiều.

Sơ đồ tổng quát của một hệ thống xử lý ảnh:



Hình 1.2: Các bước cơ bản trong một hệ thống xử lý ảnh

➤ **Thu nhận ảnh (Image acquisition)**

Các thiết bị thu nhận ảnh có hai loại chính ứng với hai loại ảnh thông dụng là Raster và Vector. Các thiết bị thu nhận ảnh Raster là camera còn các thiết bị thu nhận ảnh Vector là sensor hoặc bộ số hoá (digitalizer) hoặc được chuyển đổi từ ảnh Raster. Các thiết bị thu nhận ảnh thông thường gồm camera cộng với bộ chuyển đổi tương tự số AD (Analog to Digital) hoặc scanner chuyên dụng. Các thiết bị thu nhận ảnh này có thể cho ảnh đen trắng hoặc ảnh màu. Đầu ra của scanner là ảnh ma trận số mà ta quen gọi là bản đồ ảnh (ảnh Bitmap). Bộ số hoá (digitalizer) sẽ tạo ảnh vector có hướng. Nhìn chung, các hệ thống thu nhận ảnh thực hiện hai quá trình:

- Cảm biến: biến đổi năng lượng quang học (ánh sáng) thành năng lượng điện.
- Tổng hợp năng lượng điện thành ảnh.

➤ **Tiền xử lý (Image processing)**

Tiền xử lý là bước tăng cường ảnh để nâng cao chất lượng ảnh. Do những nguyên nhân khác nhau: có thể do chất lượng thiết bị thu nhận ảnh, do nguồn sáng hay do nhiễu, ảnh có thể bị suy biến. Do vậy cần phải tăng cường và khôi phục lại ảnh để làm nổi bật một số đặc tính chính của ảnh, hay làm cho ảnh gần giống nhất với trạng thái gốc - trạng thái trước khi ảnh bị biến dạng.

➤ **Trích chọn đặc điểm (Feature extraction)**

Vì lượng thông tin chứa trong ảnh là rất lớn, trong khi đó đa số ứng dụng chỉ cần một số thông tin đặc trưng nào đó, cần có bước trích chọn đặc điểm để giảm lượng thông tin không lờ ấy. Các đặc trưng của ảnh thường gồm: mật độ xám, phân bố xác suất, phân bố không gian, biên ảnh.

➤ **Hậu xử lý**

Nếu lưu trữ ảnh trực tiếp từ các ảnh thô (brut image) theo kiểu bản đồ ảnh đòi hỏi dung lượng bộ nhớ lớn, tốn kém mà nhiều khi không hiệu quả theo quan điểm ứng dụng. Thường người ta không biểu diễn toàn bộ ảnh thô mà tập trung đặc tả các đặc trưng của ảnh như biên ảnh (boundary) hay vùng ảnh (region) . Một số phương pháp biểu diễn thường dùng:

- Biểu diễn mã loạt dài (Run-Length Code).
- Biểu diễn mã xích (Chaine -Code).

- Biểu diễn mã tứ phân (Quad-Tree Code).

Ảnh là một đối tượng khá phức tạp về đường nét, độ sáng tối, dung lượng điểm ảnh, môi trường để thu ảnh phong phú kéo theo nhiều. Trong nhiều khâu xử lý và phân tích ảnh ngoài việc đơn giản hóa các phương pháp toán học đảm bảo tiện lợi cho xử lý, người ta mong muốn bắt chước quy trình tiếp nhận và xử lý ảnh theo cách của con người. Trong các bước xử lý đó, nhiều khâu hiện nay đã xử lý theo các phương pháp trí tuệ con người. Vì vậy, ở đây các cơ sở tri thức- hệ quyết định được phát huy.

➤ **Đôi sánh rút ra kết luận**

So sánh ảnh sau bước hậu xử lý với mẫu chuẩn hoặc ảnh đã được lưu trữ từ trước, phục vụ cho các mục đích nhận dạng và nội suy ảnh.

1.1.2 Các vấn đề cơ bản trong xử lý ảnh

1.1.2.1 Một số khái niệm

a. Ảnh

Trong thực tế ảnh liên tục về không gian và độ sáng. Để xử lý được bằng máy tính thì ảnh cần được số hóa. Số hóa ảnh là sự biến đổi gần đúng một ảnh liên tục thành tập điểm phù hợp với ảnh thật về vị trí và độ sáng. Khoảng cách giữa các điểm ảnh đó được thiết lập sao cho mắt người không phân biệt được ranh giới giữa chúng. Mỗi điểm như vậy gọi là điểm ảnh (pixel).

Từ đó ta có định nghĩa : Điểm ảnh được xem như là dấu hiệu hay cường độ sáng tại 1 toạ độ trong không gian của đối tượng và ảnh được xem như là một tập hợp các điểm ảnh. Khi được số hoá, nó thường được biểu diễn bởi bảng hai chiều $I(n,p)$: n dòng và p cột. Ta nói ảnh gồm n x p điểm ảnh. Người ta thường kí hiệu $I(x,y)$ để chỉ một điểm ảnh. Thường giá trị của n chọn bằng p và bằng 256. Một điểm ảnh có thể lưu trữ trên 1, 4, 8 hay 24 bit .

Về mặt toán học có thể xem ảnh là một hàm hai biến $f(x,y)$ với x, y là các biến toạ độ. Giá trị số ở điểm (x,y) tương ứng với giá trị xám hoặc độ sáng của ảnh (x là các cột, y là các hàng). Giá trị của hàm ảnh $f(x,y)$ được hạn chế trong phạm vi của các số nguyên dương: $0 \leq f(x,y) \leq f_{max}$. Thông thường đối với ảnh xám, giá trị f_{max} là 255 ($2^8=256$) bởi vì mỗi phần tử ảnh được mã hóa bởi một byte. Khi quan tâm đến ảnh

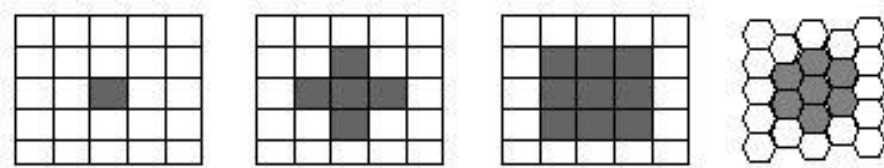
màu, ta có thể mô tả màu qua ba hàm số: $R(x,y)$ của màu đỏ, $G(x,y)$ của màu xanh lục và $B(x,y)$ của màu xanh lam.

Ảnh có thể được biểu diễn theo mô hình Vector hoặc mô hình Raster:

- **Mô hình Raster**

Đây là mô hình biểu diễn ảnh thông dụng nhất hiện nay. Ảnh được biểu diễn dưới dạng ma trận các điểm ảnh. Tùy theo nhu cầu thực tế mà mỗi điểm ảnh có thể được biểu diễn bởi một hay nhiều bit. Mô hình Raster rất thuận lợi cho hiển thị và in ấn.

Khi xử lý các ảnh Raster, chúng ta quan tâm đến mối quan hệ trong vùng lân cận của các điểm ảnh. Các điểm ảnh có thể xếp hàng trên một lưới (Raster) hình vuông, lưới hình lục giác hoặc theo một cách hoàn toàn ngẫu nhiên với nhau:



Hình 1.3: Quan hệ giữa các điểm ảnh

- **Mô hình Vector**

Biểu diễn ảnh ngoài mục đích tiết kiệm không gian lưu trữ, dễ dàng cho hiển thị và in ấn, còn phải đảm bảo dễ dàng trong lựa chọn, sao chép, di chuyển, tìm kiếm... Theo những yêu cầu này, kỹ thuật biểu diễn Vector tỏ ra ưu việt hơn. Trong mô hình Vector người ta sử dụng hướng giữa các Vector của điểm ảnh lân cận để mã hoá và tái tạo hình ảnh ban đầu. Ảnh Vector được thu nhận trực tiếp từ các thiết bị số hóa như Digital hoặc được chuyển đổi từ ảnh Raster thông qua các chương trình số hóa. Công nghệ phần cứng cung cấp những thiết bị xử lý với tốc độ nhanh và chất lượng cao cho cả đầu vào và ra, nhưng lại chỉ hỗ trợ cho ảnh Raster. Do vậy, những nghiên cứu về biểu diễn Vector đều tập trung chuyển đổi từ ảnh Raster.

b. Độ phân giải

Độ phân giải là số lượng điểm ảnh (pixel) dùng để tập hợp thành hình ảnh. Số lượng điểm ảnh càng nhiều và càng nhỏ thì độ nét và chi tiết ảnh sẽ càng cao.

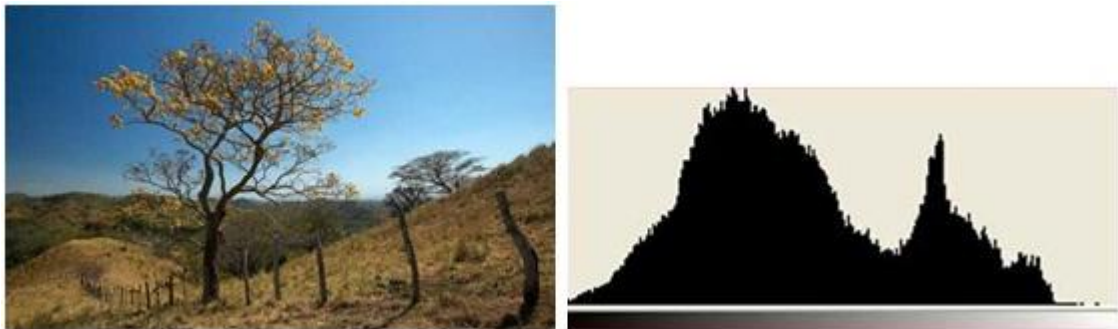
Có 3 cách để biểu thị độ phân giải ảnh:

- Biểu thị bằng số lượng điểm ảnh theo chiều dọc và chiều ngang của ảnh (ví dụ: 1024 x 768)
- Biểu thị bằng tổng số điểm ảnh trên 1 tấm ảnh (960.000 pixel)
- Biểu thị bằng số lượng điểm ảnh có trên 1 inch (ppi) hoặc số chấm(dot) có trên 1 inch (dpi)

c. *Mức xám của ảnh*

Mức xám là kết quả sự mã hoá tương ứng một cường độ sáng của mỗi điểm ảnh với một giá trị số - kết quả của quá trình lượng hoá. Cách mã hoá kinh điển thường dùng 16, 32 hay 64 mức. Mã hoá 256 mức là phổ dụng nhất do lý do kỹ thuật. Vì $2^8 = 256$ (0, 1, ..., 255), nên với 256 mức, mỗi pixel sẽ được mã hoá bởi 8 bit.

Lược đồ xám hay còn gọi là *biểu đồ tần suất* được biểu diễn trong hệ tọa độ vuông góc Oxy. Trong hệ tọa độ này, trục hoành biểu diễn cho số mức xám từ 0 đến N, N là số mức xám (thường xét với mức 256). Trục tung biểu diễn số điểm ảnh cho một mức xám (số điểm ảnh có cùng mức xám). Cũng có thể biểu diễn là: trục tung là tỉ lệ số điểm ảnh có cùng mức xám trên tổng số điểm ảnh.

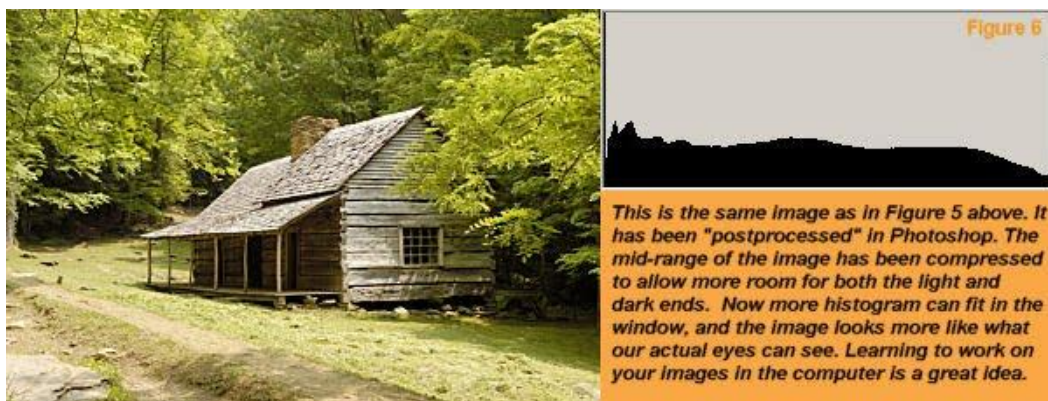


Hình 1.4: Lược đồ xám của ảnh

Lược đồ xám cung cấp rất nhiều thông tin về phân bố mức xám của ảnh. Theo thuật ngữ của xử lý ảnh gọi là tính động của ảnh. Tính động của ảnh cho phép phân tích trong khoảng nào đó phân bố phần lớn các mức xám của ảnh: ảnh rất sáng hay ảnh rất đậm. Nếu ảnh sáng, lược đồ xám nằm bên phải (mức xám cao), còn ảnh đậm lược đồ xám nằm bên trái (mức xám thấp). Hình 1.5 và 1.6 là một ví dụ :



Hình 1.5: Bức ảnh bị dư tối có lược đồ xám tập trung nhiều bên trái



Hình 1.6: Hình ảnh sau khi được chỉnh sửa lược đồ xám đã được trải đều

1.1.2.2 Ứng dụng của xử lý ảnh

Ban đầu, các kỹ thuật xử lý ảnh đây chủ yếu được sử dụng để nâng cao chất lượng hình ảnh, chính xác hơn là tạo cảm giác về sự gia tăng chất lượng ảnh quang học trong mắt người quan sát. Thời gian gần đây, phạm vi ứng dụng xử lý ảnh mở rộng không ngừng, có thể nói hiện không có lĩnh vực khoa học nào không sử dụng các thành tựu của công nghệ xử lý ảnh số.

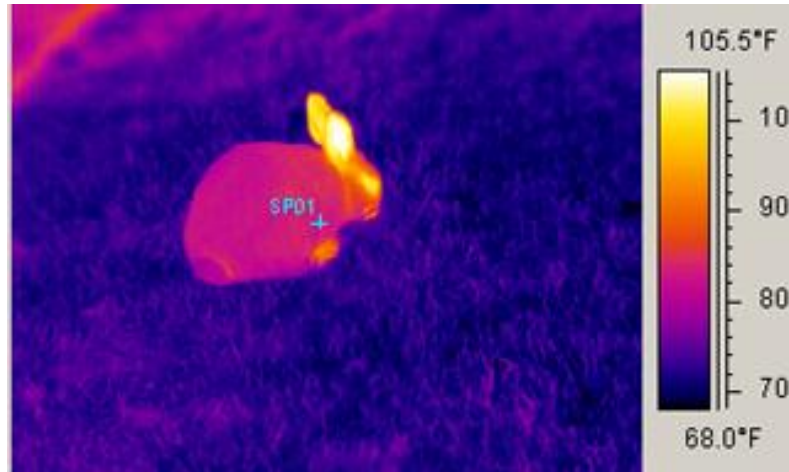
Trong y học các thuật toán xử lý ảnh cho phép biến đổi hình ảnh được tạo ra từ nguồn bức xạ X-ray hay nguồn bức xạ siêu âm thành hình ảnh quang học trên bề mặt film x-quang hoặc trực tiếp trên bề mặt màn hình hiển thị. Hình ảnh các cơ quan chức năng của con người sau đó có thể được xử lý tiếp để nâng cao độ tương phản, lọc, tách các thành phần cần thiết (chụp cắt lớp) hoặc tạo ra hình ảnh trong không gian ba chiều (siêu âm 3 chiều).

Trong lĩnh vực địa chất, hình ảnh nhận được từ vệ tinh có thể được phân tích để xác định cấu trúc bề mặt trái đất. Kỹ thuật làm nổi đường biên (image enhancement) và khôi phục hình ảnh (image restoration) cho phép nâng cao chất lượng ảnh vệ tinh và tạo ra các bản đồ địa hình 3-D với độ chính xác cao.



Hình 1.7: Bản đồ Việt Nam nhận được từ vệ tinh

Trong ngành khí tượng học, ảnh nhận được từ hệ thống vệ tinh theo dõi thời tiết cũng được xử lý, nâng cao chất lượng và ghép hình để tạo ra ảnh bề mặt trái đất trên một vùng rộng lớn, qua đó có thể thực hiện việc dự báo thời tiết một cách chính xác hơn. Dựa trên các kết quả phân tích ảnh vệ tinh tại các khu vực đông dân cư còn có thể dự đoán quá trình tăng trưởng dân số, tốc độ ô nhiễm môi trường cũng như các yếu tố ảnh hưởng tới môi trường sinh thái. Ảnh chụp từ vệ tinh có thể thu được thông qua các thiết bị ghi hình cảm nhận được tia sáng quang học (450-520 nm), hoặc tia hồng ngoại (760-900 nm). Thiết bị thu hình nhạy cảm với vật thể bức xạ các tia trong miền hồng ngoại sẽ cho ra những bức ảnh trong đó vật thể có nhiệt độ thấp sẽ được phân biệt rõ ràng so với vật thể có nhiệt độ cao hơn. Như vậy việc lựa chọn các thiết bị ghi hình khác nhau sẽ tạo ra ảnh có đặc tính khác nhau, tùy thuộc vào mục đích sử dụng trong các lĩnh vực khoa học cụ thể.



Hình 1.8: Ảnh hồng ngoại

Xử lý ảnh còn được sử dụng nhiều trong các hệ thống quản lý chất lượng và số lượng hàng hóa trong các dây chuyền tự động, ví dụ như hệ thống phân tích ảnh để phát hiện bọt khí bên vật thể đúc bằng nhựa, phát hiện các linh kiện không đạt tiêu chuẩn (bị biến dạng) trong quá trình sản xuất hoặc hệ thống đếm sản phẩm thông qua hình ảnh nhận được từ camera quan sát. Xử lý ảnh còn được sử dụng rộng rãi trong lĩnh vực hình sự và các hệ thống bảo mật hoặc kiểm soát truy cập: quá trình xử lý ảnh với mục đích nhận dạng vân tay hay khuôn mặt cho phép phát hiện nhanh các đối tượng nghi vấn cũng như nâng cao hiệu quả hệ thống bảo mật cá nhân cũng như kiểm soát ra vào. Ngoài ra, có thể kể đến các ứng dụng quan trọng khác của kỹ thuật xử lý ảnh tĩnh cũng như ảnh động trong đời sống như tự động nhận dạng, nhận dạng mục tiêu quân sự, máy nhìn công nghiệp trong các hệ thống điều khiển tự động, nén ảnh tĩnh, ảnh động để lưu và truyền trong mạng viễn thông v.v.

1.2 Khảm ảnh

1.2.1 Khái niệm khảm ảnh

Khảm (mosaic) là nghệ thuật ghép mảnh từ các vật liệu như đá, gạch, thủy tinh, ... có từ thời La Mã cổ đại và được ứng dụng nhiều trong kiến trúc công cộng như nhà thờ, đền đài, cung điện... Ngày nay, mosaic được ứng dụng rộng rãi trong nhiều lĩnh vực, đặc biệt nhất là trang trí nội ngoại thất kiến trúc. Tuy nhiên không chỉ vậy mà khảm còn được ứng dụng vào trong xử lý ảnh để tạo ra các bức hình mang phong cách mới mẻ nhằm phục vụ cho rất nhiều mục đích khác nhau từ trưng bày trong các phòng triển lãm đến việc tạo các hình ảnh quảng cáo độc đáo.

Khảm ảnh là việc tạo ra hình ảnh mới bằng cách ghép các ảnh nhỏ vào 1 ảnh lớn sao cho khi nhìn tổng thể vào ảnh lớn thì ta vẫn có thể nhìn thấy nội dung của bức ảnh lớn trước đó. Khảm ảnh có thể được chia thành hai loại chính đó là: khảm ảnh toàn cảnh và khảm ảnh nhiều lớp.



Hình 1.9: Ảnh con chim được khảm từ rất nhiều ảnh nhỏ

- ***Khảm toàn cảnh***

Ảnh toàn cảnh (Panorama), Panorama bắt nguồn từ tiếng Hy Lạp, có nghĩa là góc nhìn rộng trong một không gian nhất định. Kỹ thuật Panorama được sử dụng trong rất nhiều môn nghệ thuật như sơn, vẽ, dựng hình 3D và đặc biệt là trong nhiếp ảnh cho bạn một bức ảnh toàn cảnh tuyệt vời. Một bức ảnh thông thường chỉ được chụp với một góc 90 độ, nên rất khó để bạn có thể thu lại hết toàn cảnh không gian rộng lớn mà bạn mong muốn. Với Panorama, một bức ảnh hiển thị cả 360⁰ là một chuyện có thể thực hiện dễ dàng.

Ngày xưa, để tạo ra một bức ảnh toàn cảnh Panorama, các nhà nhiếp ảnh chuyên nghiệp đã phải sử dụng những máy ảnh đắt tiền, và ngồi nhiều ngày liền trong phòng tối ráp những bức ảnh lại với nhau bằng kỹ thuật phơi sáng phức tạp.

Nhưng ngày nay, trong thời đại kỹ thuật số, công việc này trở nên nhẹ nhàng hơn rất nhiều. Bạn có thể sắm cho mình dòng máy ảnh chuyên nghiệp DSLR chuyên chụp ảnh Panorama nhưng có một cách không cần đến loại máy ảnh chuyên dụng đó, chỉ cần loại máy ảnh kỹ thuật số thông thường cũng đã hỗ trợ tính năng này tuy nhiên vẫn còn cách khác đó là sử dụng máy ảnh thông thường chụp các bức hình liên tiếp sau

đó đưa lên máy tính và sử dụng các phần mềm ghép hình là bạn đã có một tấm hình Panorama độc đáo. Ảnh Panorama được chia làm nhiều loại, nhưng phổ biến nhất là hai kiểu chụp sau:

Thứ nhất, chụp một cảnh ra thành nhiều file, sau đó dùng phần mềm ghép chúng lại với nhau. Với cách chụp này, bạn có thể chụp phong cảnh theo chiều ngang (chiều rộng lớn gấp nhiều lần chiều dài để lấy hết toàn bộ khung cảnh rộng lớn), hoặc chụp theo chiều dọc (còn được gọi là vertorama, chiều dài lớn gấp nhiều lần chiều rộng để lấy hết chiều cao những tòa nhà cao tầng).



Hình 1.10: Ảnh Panorama được ghép từ 4 hình ảnh chụp liên tiếp

Thứ hai là Polar Panorama. Chụp 360 độ bằng cách dùng chân đờ giáp vòng, chụp cả trên lẫn dưới theo dạng hình cầu (biến không gian lớn thành một hành tinh nhỏ).



Hình 1.11: Ảnh polar panorama

Một số phần mềm hỗ trợ cho việc khâu, ghép ảnh toàn cảnh là : Photoshop, Panorama Factory, AutoStitch, Hugin, Panorama Maker, Pixtra OmniStitcher ...

Hầu hết các phần mềm ghép ảnh toàn cảnh đều qua 3 công đoạn chính như sau:

Bước 1: Sắp xếp các ảnh đầu vào theo thứ tự thích hợp

Bước 2: Phần mềm sẽ tính toán và ghép các ảnh đó với nhau. Các ảnh sẽ được xếp trùng một phần lên nhau và các phần ảnh chung sẽ được ghép lại một cách hợp lý. Trong khâu này còn một phần cũng khá quan trọng đó là nắn chỉnh các đường biên trong ảnh. Bởi ảnh ban đầu chụp thường không thể hoàn toàn trùng khít nhau do góc độ chụp, ánh sáng, hay do thiết bị chụp ... vì vậy cần phải nắn chỉnh lại các đường biên sao cho chúng hợp lý nhất giữa các ảnh.

Bước 3: Sau khi đã ghép xong thì vấn đề đặt ra là phải làm trơn các vùng chuyển tiếp giữa các ảnh để bức ảnh thể hiện được không gian liên tục, đồng nhất. Phần mềm sẽ sử dụng các kỹ thuật khác nhau để sao cho

bức ảnh đẹp nhất. Một số kỹ thuật thường thấy đó là khớp biểu đồ tần suất, các kỹ thuật lọc ...

Tuy nhiên việc khảm ảnh toàn cảnh còn thể hiện ở góc độ cao hơn là ghép toàn cảnh từ nhiều file khác nhau theo nhiều hướng của bức ảnh chứ không chỉ ghép theo chiều ngang hoặc chiều dọc.



Hình 1.12: Ảnh khảm toàn cảnh

- **Khảm nhiều lớp**

Trong lĩnh vực hội họa, để tạo ra một bức tranh khảm thì họa sĩ cần phải hình dung, tưởng tượng ra bức tranh toàn cảnh mà mình đang vẽ đồng thời vẽ các chi tiết nhỏ bên trong để cuối cùng cho ra một bức tranh khảm. Điều này đòi hỏi rất cao về mặt chuyên môn và năng khiếu thẩm mỹ đồng thời tiêu tốn rất nhiều thời gian và công sức. Trong xử lý ảnh cũng tạo ra được những bức tranh khảm như vậy nhưng có điều nó làm hơi ngược một chút. Từ bức tranh tổng thể ban đầu, bằng các kỹ thuật xử lý khác nhau thì các bức ảnh nhỏ được lồng ghép vào đó tạo nên bức ảnh mới. Tất nhiên là nếu nhìn một cách tổng thể thì nó vẫn chính là bức tranh lớn ban đầu có điều nó khác đi một chút bởi những chi tiết bên trong đã được thay thế bởi các hình ảnh đơn lẻ.

Về quy trình tạo nên một hình ảnh khảm nhiều lớp này thì việc đầu tiên không thể thiếu đó chính là 2 thành phần chính: ảnh nguồn (ảnh dùng làm nền toàn cảnh) và

ảnh mẫu (các ảnh nhỏ được dùng để ghép vào ảnh mục tiêu các ảnh này được thu thập càng đa dạng càng tốt và được lưu chung tại một thư mục).

Có rất nhiều phương pháp khác nhau để thực hiện khâu ảnh, sau đây là một số phương pháp thường được sử dụng:

Thứ nhất theo như trang <http://scien.stanford.edu> thì công đoạn tạo ra ảnh khâu gồm 5 bước (Giải thuật này được dựa trên giải thuật của giáo sư Walldel):

Bước 1: Xử lý các ảnh nhỏ:

- Chuyển các ảnh nhỏ này thành ảnh đen trắng. Thực hiện bằng cách lấy trung bình các giá trị R, G, B rồi đặt đồng thời các giá trị R, G, B đó bằng giá trị trung bình.
- Thu nhỏ các hình này lại nếu cần thiết. Chính kích thước cho các ảnh nhỏ đồng nhất nhau.



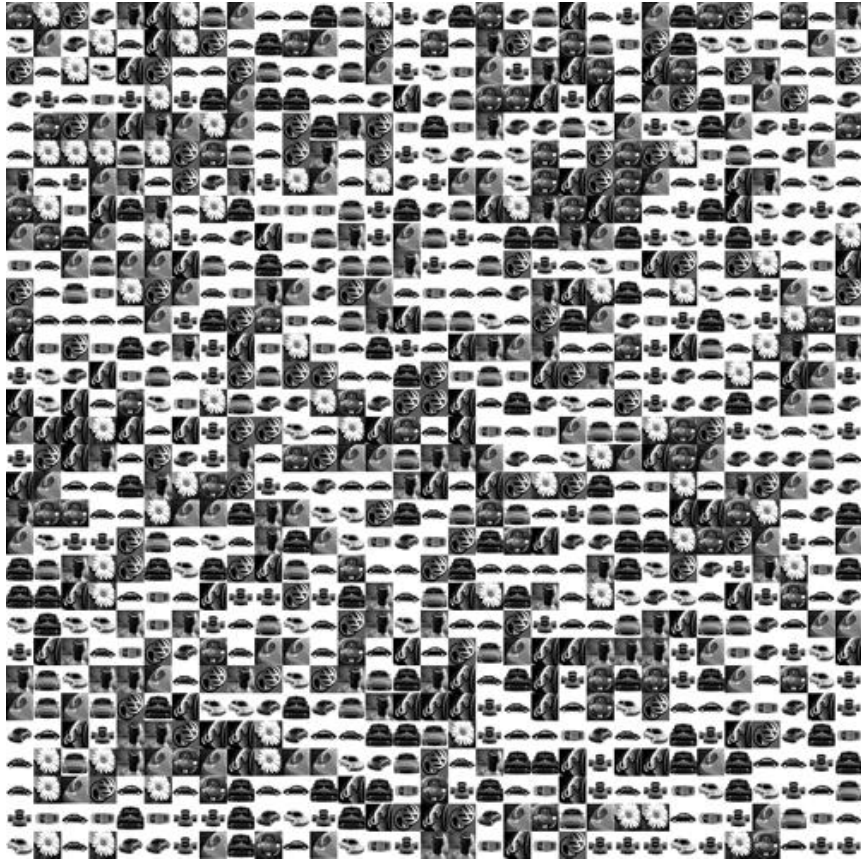
Hình 1.13 Ảnh mẫu sau khi được xử lý

Bước 2: Thay đổi kích thước ảnh gốc

- Để có tạo ra các lưới trên ảnh gốc bằng với kích thước các ảnh mẫu một cách nhanh chóng và dễ dàng thì ảnh gốc sẽ được cắt để kích thước của nó là một bội số của kích thước của các ảnh mẫu nhỏ (đã được xử lý).

Bước 3: Tạo mẫu khâu

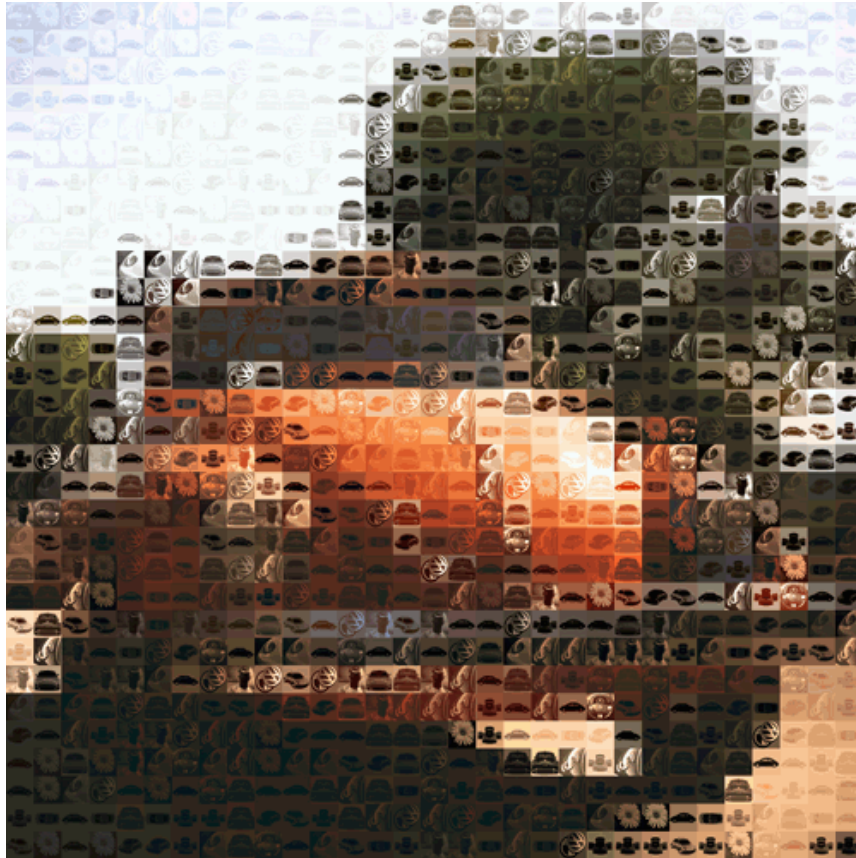
- Tạo một ảnh khâu là khái niệm gần gũi hơn với việc xây dựng một hình ảnh mới dựa trên đặc điểm của hình ảnh gốc, thay vì làm thay đổi hình ảnh cơ bản, hoặc nhúng các hình ảnh nhỏ vào nó, vv.
- Để tạo một mẫu khâu thì thực hiện các bước như sau:
 - Tạo mới một hình ảnh trống có kích thước bằng với kích thước ảnh gốc đã được cắt.
 - Phủ đầy lên ảnh trống này một cách tùy ý các ảnh mẫu nhỏ theo một lưới được dựng sẵn (mắt lưới này có kích thước bằng các ảnh nhỏ)



Hình 1.14: Mẫu ảnh khảm

Bước 4: Hòa trộn (blend) ảnh khảm mẫu với ảnh gốc

- Tại bước này, chương trình sẽ dùng thuật toán để phân tích mỗi vùng trên các lưới của ảnh khảm mẫu với ảnh gốc để tập hợp đặc tính của cả 2 vùng. Đó chính là giá trị R, G, B và phân tích giá trị suy biến.
- Sử dụng các giá trị đó để chuyển giá trị R, G, B tại mỗi vùng trên ảnh khảm mẫu thành chỉ số màu sắc được lấy từ các vùng tương ứng trên ảnh gốc.



Hình 1.15: Ảnh khảm được tạo ra bằng phương pháp trên

Khác với giải thuật của giáo sư Walldel. Giải thuật cũng như thuật toán tạo ảnh khảm của Robert Silvers thì khác. Cũng sử dụng các ảnh nhỏ nhưng số lượng ảnh nhỏ này rất lớn, cũng không cần chuyển các ảnh nhỏ này về đen trắng. Ông chia ảnh gốc theo các lưới nhỏ có kích thước bằng các ảnh nhỏ rồi so sánh các vùng đó với tất cả các ảnh nhỏ trong tập mẫu rồi sẽ lấy ảnh mẫu có thông số gần với vùng đó nhất để thay thế vào ảnh gốc.

Ngoài ra trên trang <http://www.picturemosaics.com> cũng đề cập tới một phương pháp khảm ảnh rất hay, nó cho ra những bức tranh khảm với chất lượng rất tốt. Ở đây không giống như các phần mềm khác là chia bức ảnh nguồn thành các lưới với các ô hình vuông hoặc hình chữ nhật cố định mà họ đã chia ảnh nguồn thành các “vùng mềm” (soft zones) mà họ còn gọi là vùng thích nghi (Adaptive Regions™). Điều này sẽ giúp cho một số vùng quan trọng trong bức ảnh nguồn của bạn được quan tâm đặc biệt như vậy nội dung chính của bức ảnh sẽ được bảo toàn nhất. Bằng việc sử dụng kỹ thuật kép để phân tích các hình ảnh đối với các khuôn dạng mà màu sắc họ đã cho thấy hiệu quả của việc sử dụng các vùng thích nghi để tối ưu cho việc lựa chọn các vùng

xuất hiện trên bức ảnh. Họ xác định và tối ưu các thành phần quan trọng trong bức ảnh như văn bản, khuôn mặt, logo, .v.v.. Và tất cả những ô được chia trên ảnh gốc đều rất nhỏ, điều này khiến cho bức ảnh càng giống với ảnh gốc.



Hình 1.16: Ảnh gốc



Hình 1.17: Làm nổi các màu sắc chủ đạo



Hình 1.18: Khám các thành phần chủ đạo



Hình 1.19: Ảnh hoàn thiện

Ngày nay, có rất nhiều các phần mềm khám ảnh đã và đang được xây dựng. Trong đó phải kể đến những cái tên như : Photoshop, Easy Mosaic, Mazaika, AndreaMosaic, Photomosaic Generator ...

1.2.2 Các kỹ thuật chính được dùng trong khảm ảnh

- Đan đa phân giải (Multiresolution spline).
- Khớp biểu đồ tần suất (Histogram matching).
- Khớp cạnh (Edge matching).
- Phân tích cấu trúc (Texture analysis).
- Phân tích hình khối (Shape analysis).
- Trích trộn đặc trưng.
- S.M.A.R.T.TM[7]

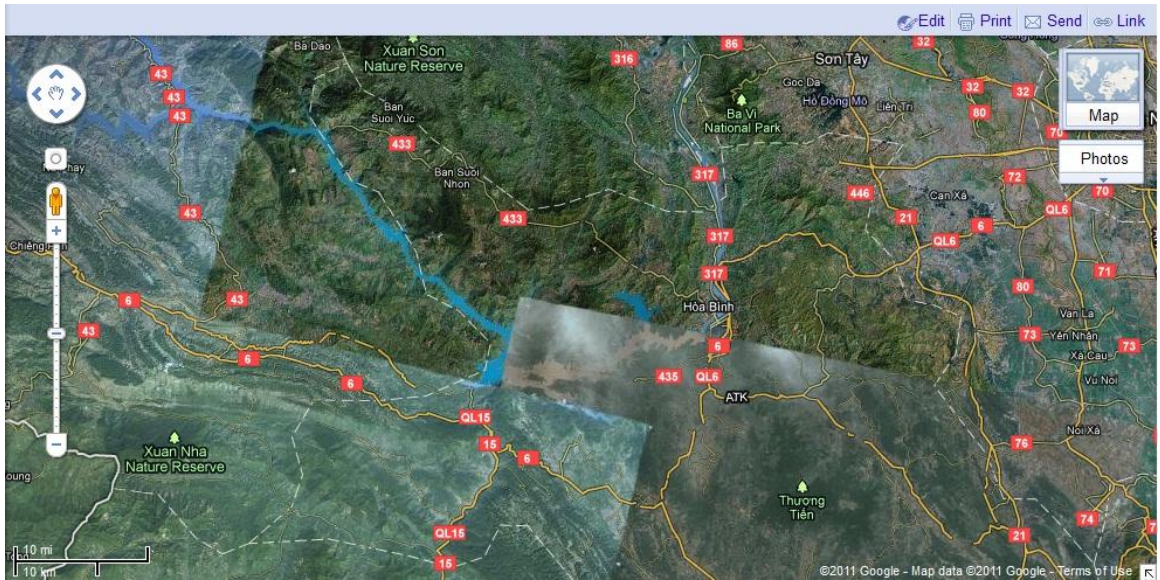
Ngoài các kỹ thuật chính trên thì các kỹ thuật lọc cũng được áp dụng để làm mịn đi các vùng chuyển tiếp giữa các ảnh ghép làm cho bức ảnh thành phẩm mịn hơn.

1.2.3 Ứng dụng của khảm ảnh

Không chỉ dừng lại ở việc tạo ra các hình ảnh độc đáo mang phong cách mới lạ, khảm ảnh còn được ứng dụng rất rộng rãi trong nhiều lĩnh vực khác nhau.

1.2.3.1 Bản đồ số

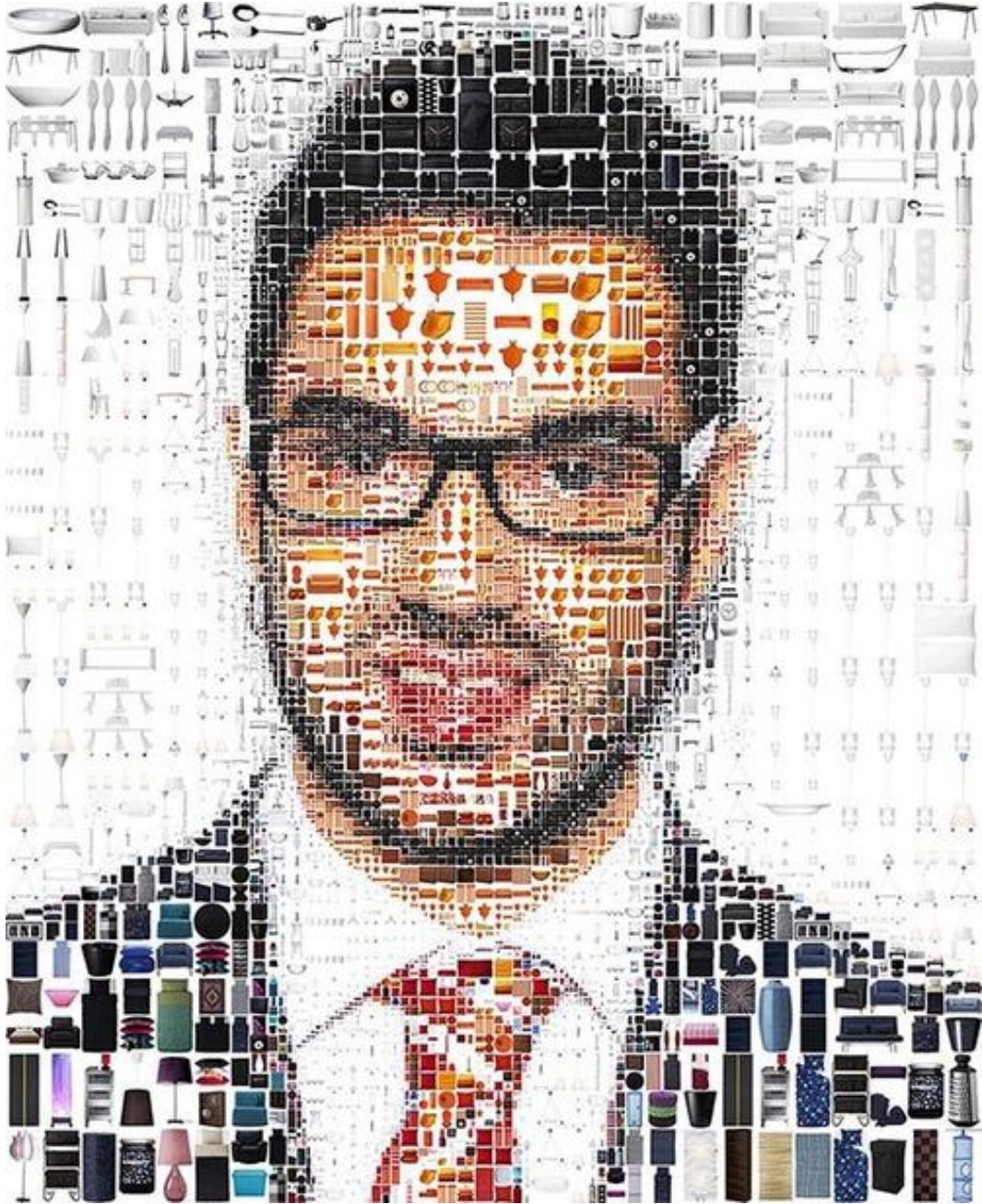
Chắc hẳn cái tên bản đồ số không còn xa lạ gì đối với những người dùng internet ngày nay. Hoặc kể đến cái tên nổi bật đó chính là google maps, vietbando. Họ cho phép người dùng xem bản đồ ở chế độ vệ tinh. Khi xem bản đồ này tất cả mọi người đều nhận thấy một điều rất rõ đó chính là bản đồ được ghép lại từ rất nhiều hình ảnh khác nhau, điều này được nhận thấy rất rõ qua màu sắc có phần khác biệt giữa từng vùng bản đồ, hay thậm chí là hình ảnh không khớp của các đám mây, con đường ... Bản đồ số chính là một ví dụ điển hình của kỹ thuật khảm ảnh. Những hình ảnh này được chụp từ các ống kính có độ phân giải rất cao trên vệ tinh, ảnh được gửi về các trung tâm nghiên cứu quốc gia rồi từ đây, các hình ảnh đã được ghép lại với nhau tạo nên bản đồ số hoàn chỉnh.



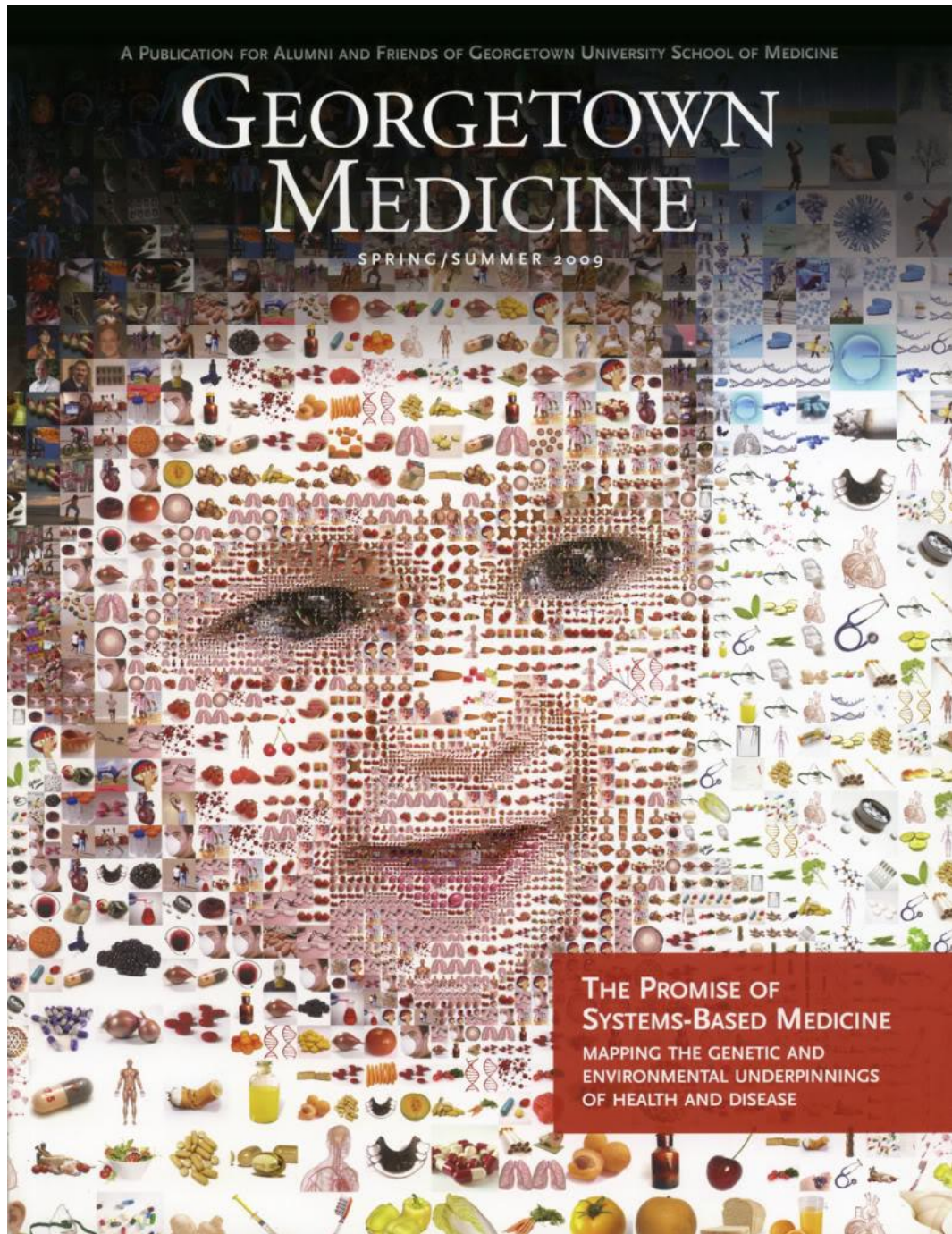
Hình 1.20: Ảnh chụp từ google maps

1.2.3.2 Truyền thông quảng cáo

Giờ đây, mọi người đã quá quen thuộc với các bức ảnh thông thường, nên nó ít gây được sự chú ý. Chính vì thế mà một bức ảnh mang phong cách khám mới lạ chắc chắn sẽ thu hút được sự chú ý của mọi người hơn. Điều này có tác dụng tích cực trong các lĩnh vực quảng cáo, giới thiệu sản phẩm. Ví dụ như ảnh quảng cáo sản phẩm của hãng Ikea – một hãng sản xuất đồ gia dụng, ảnh bìa tạp chí y học Georgetown ...



Hình 1.21: Ảnh quảng cáo của hãng IKEA



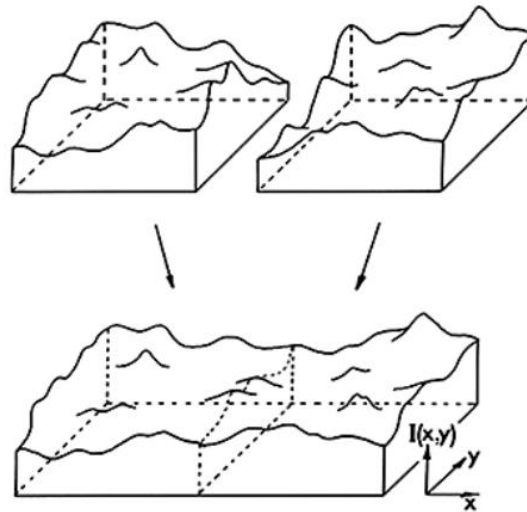
Hình 1.22: Ảnh bìa tạp chí y học Georgetown

CHƯƠNG 2: KỸ THUẬT KHẢM ẢNH

2.1 Kỹ thuật đan đa phân giải (Multiresolution spline)

Kỹ thuật đan đa phân giải là kỹ thuật được nghiên cứu bởi PETER J. BURT và EDWARD H. ADELSON thuộc trung tâm nghiên cứu RCA David Sarnoff.

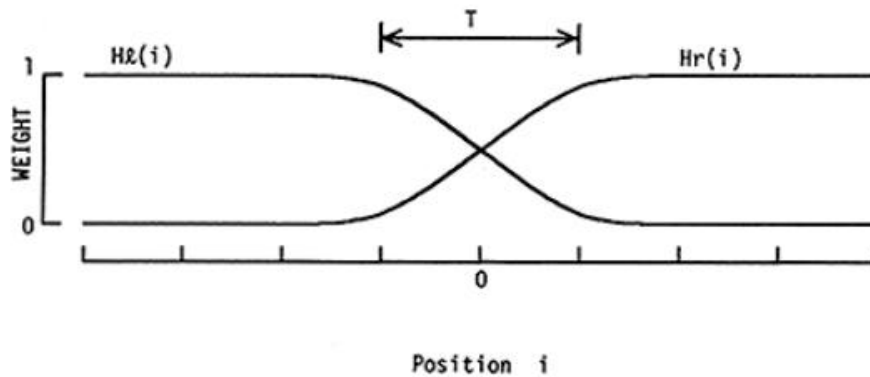
Kỹ thuật này được dùng để kết hợp hai hoặc nhiều ảnh vào một ảnh khảm lớn hơn. Đầu tiên, các ảnh dùng để đan sẽ được phân tích thành tập band-pass[5] các thành phần hình ảnh đã được lọc. Sau đó, các ảnh thành phần sẽ được ghép lại thành ảnh khảm band-pass tương ứng. Trong bước này, các ảnh thành phần được nối với nhau bằng hàm trung bình trọng số bên trong miền chuyển tiếp đó là tỉ lệ kích thước độ dài sóng được biểu diễn trong các band đó. Cuối cùng, các ảnh khảm band-pass này được tổng hợp lại thành ảnh khảm mong muốn. Bằng cách này hàm nối được làm khớp với tỉ lệ với các đường bao bên trong những ảnh đó. Khi các đường bao thô xuất hiện gần biên, chúng được pha trộn dần dần qua một khoảng cách tương đối lớn mà không được làm mờ đi các chi tiết xung quanh đường viền.



Hình 2.1: Hai hình ảnh được nối ghép sao cho đường nối càng mịn càng tốt

Đối với các ảnh chụp từ kính viễn vọng, trong trường hợp này thì kỹ thuật khảm ảnh được dùng để tạo ra cả hình ảnh có phạm vi rộng lớn hoặc là mức độ chi tiết hơn một ảnh đơn. Trong lĩnh vực quảng cáo hay đồ họa máy tính, công nghệ này có thể tổng hợp lên hình ảnh từ các thành phần khác nhau.

Một vấn đề chung cho tất cả các ứng dụng của khảm ảnh đó là khi các ảnh được ghép nối với nhau thì đường biên giữa chúng phải được làm ẩn đi. Chỉ cần một khác biệt nhỏ trong mức xám của ảnh trên đường biên rộng cũng rất dễ nhận thấy. Thật không may là điều này rất khó tránh khỏi, có thể do vị trí đặt máy hay trong khâu xử lý. Như vậy một kỹ thuật cần đòi hỏi đó sẽ thay đổi mức xám của hình ảnh trong vùng lân cận của đường biên để đạt được sự chuyển tiếp mượt giữa các ảnh. Hai hình ảnh được dùng để ghép nối có thể xem như hai bề mặt, nơi mà cường độ hình ảnh $I(x, y)$ tương ứng được chiếu lên không gian x, y . Vấn đề đặt ra, như trong hình minh họa 2.1, có thể được nói như sau: Làm thế nào để hai mặt đó có thể được làm biến dạng một cách nhẹ nhàng để có được một đường nối trơn tru? Họ đã sử dụng thuật ngữ đan ghép ảnh - “image spline” để chỉ kỹ thuật làm nên điều này. Một kỹ thuật đan ghép ảnh tốt sẽ cho ra những bức ảnh ghép có đường biên mịn.



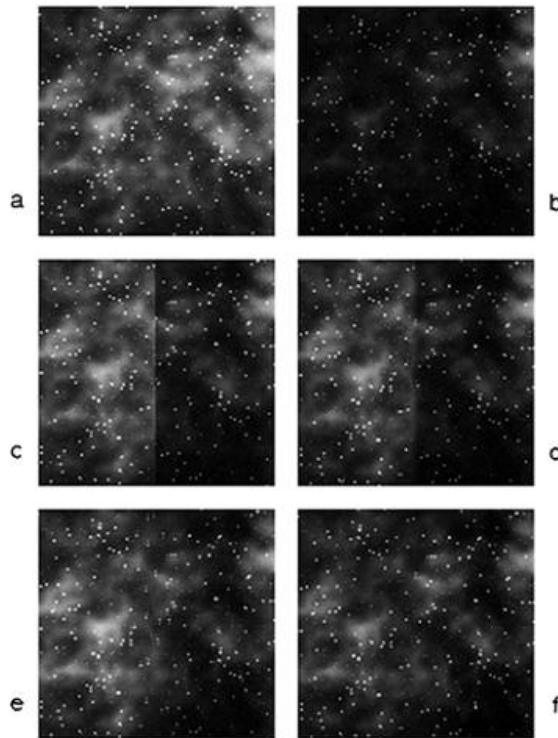
Hình 2.2: Các hàm trung bình trọng số và chiều rộng T của miền chuyển tiếp

Sự khác biệt của độ lớn trong mức xám của đường biên ảnh khảm có thể được giảm xuống mức nào đó bởi lựa chọn đúng đắn của vị trí đường biên khi đan các ảnh vào nhau. Đường nối có thể được cải thiện bằng cách thêm vào đó một đường nối tuyến tính để giá trị điểm ảnh hai bên cân bằng nhau. Một quá trình chuyển đổi vẫn mượt mà có thể thu được bằng cách sử dụng một kỹ thuật gần đây đã được đề xuất bởi Peleg [5]. Chức năng hiệu chỉnh “mịn nhất có thể” đã được xây dựng, nó thêm vào mỗi ảnh của ảnh khảm để loại bỏ sự khác biệt của đường biên. Tuy nhiên kỹ thuật này không thiết thực đối với các ảnh lớn, bởi chức năng hiệu chỉnh phải được tính toán bằng việc sử dụng một thuật toán lặp đi lặp lại. Họ quan tâm tới phương pháp đan ghép trung bình trọng số. Để bắt đầu, người ta cho rằng những hình ảnh sẽ được nối chồng lên nhau để nó có thể tính toán giá trị mức xám của các điểm trong một khu vực

chuyển tiếp như trung bình trọng số trong mỗi ảnh. Giả sử trong hai bức ảnh, $F_l(i)$ ở bên trái còn $F_r(i)$ nằm bên phải hai bức ảnh này được nối chồng lên nhau tại điểm i (biểu diễn trên một chiều để đơn giản hóa các kí hiệu). Đặt $H_l(i)$ là hàm trọng số bên trái và hàm này giảm đơn điệu từ trái sang phải và đặt $H_r(i) = 1 - H_l(i)$. Sau đó hình ảnh được đan ghép F cho bởi:

$$F(i) = H_l(i) F_l(i) + H_r(i) F_r(i) \quad (2.1)$$

Rõ ràng là với một sự lựa chọn H thích hợp, kĩ thuật trung bình trọng số sẽ cho kết quả là miền chuyển tiếp mịn. Tuy nhiên, chỉ riêng điều này chưa đảm bảo được vùng biên sẽ được làm mờ. Cho T là độ rộng của miền chuyển tiếp với H_l chạy từ 1 đến 0. Nếu T là nhỏ so với các điểm đặc trưng của ảnh thì biên có thể vẫn xuất hiện. Mặt khác, nếu T rộng so với các điểm đặc trưng của ảnh thì trong miền chuyển tiếp các chi tiết của cả hai ảnh có thể xuất hiện chồng lên nhau, giống như trong một bức ảnh được phơi sáng hai lần.



Hình 2.3: Một số thử nghiệm kĩ thuật đan ghép với hình ảnh các ngôi sao

Trong hình 2.3, các bức ảnh gốc hình 2.3a và hình 2.3b có cùng kích thước và chỉ khác nhau một số dịch chuyển không đáng kể vị trí thẳng đứng và một sự thay đổi nhỏ về mức xám. Sự khác biệt đầu tiên có thể nảy sinh từ sự biến dạng quang học hoặc độ lệch thực tế của ảnh, sự khác biệt thứ hai có thể do sự khác biệt về điều kiện khí quyển hoặc về công nghệ nhiếp ảnh.

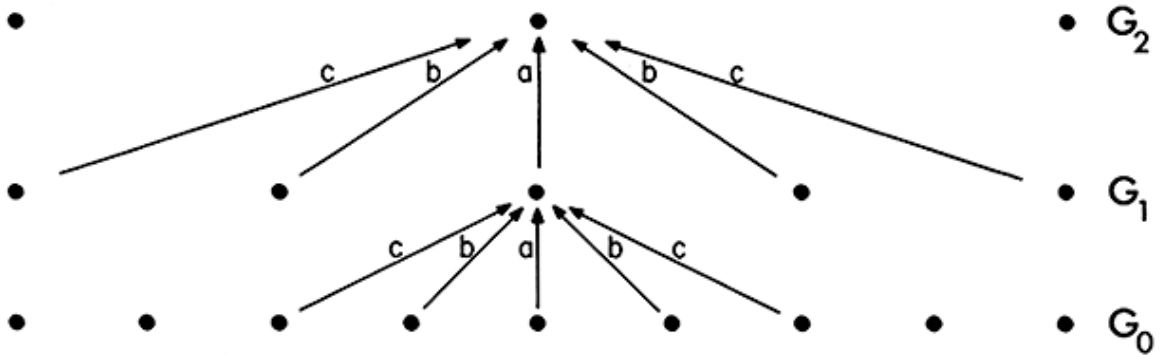
Trong ví dụ này, ảnh khảm được tạo bằng cách nối nửa trái của hình 2.3a với nửa phải của hình 2.3b. Nếu việc này được thực hiện mà không có bất kì biện pháp nào để làm mịn vùng chuyển tiếp ảnh ($T = 0$) thì đường biên sẽ xuất hiện rõ nét như trong hình 2.3c. Nếu thay vào đó, các ảnh được nối bằng phương pháp trung bình trọng số trong một miền chuyển tiếp hẹp ($T = 8$) thì đường biên đó vẫn xuất hiện nhưng nó sẽ mờ đi như trong hình 2.3d. Khi các ảnh được đan với miền chuyển tiếp rộng ($T = 64$) thì không nhìn thấy cạnh nữa nhưng trong miền chuyển tiếp các ngôi sao nhìn như được phơi sáng hai lần như hình 2.3e. Rõ ràng là kích thước vùng chuyển tiếp liên quan đến kích thước của các đặc trưng trong hình ảnh, nó đóng vai trò rất quan trọng trong việc đan ghép ảnh. Để loại bỏ đường biên hiện ra thì chiều rộng của vùng chuyển tiếp phải được so sánh ít nhất với kích thước lớn nhất của các đặc trưng trong ảnh. Mặt khác để tránh hiệu ứng phơi sáng hai lần, vùng chuyển tiếp không được lớn hơn nhiều so với đặc trưng nhỏ nhất của hình ảnh. Không có sự lựa chọn nào của T thỏa mãn đủ cả hai yêu cầu trong những hình ngôi sao của hình 2.3 bởi sự khuếch tán nền và những ngôi sao nhỏ. Các hạn chế có thể được phát biểu chính xác hơn trong ảnh không gian tần số. Đặc biệt, một T phù hợp chỉ có thể được lựa chọn nếu những hình ảnh được đan chiếm một dải tần số không gian tương đối hẹp. Với yêu cầu phức tạp, chúng ta có thể quy định rằng T sẽ được so sánh kích thước với độ lớn của tần số thấp nhất nổi bật trong ảnh. Nếu T nhỏ hơn, thì sự đan ghép sẽ giới thiệu một khía cạnh đáng chú ý khác. Mặt khác để tránh hiệu ứng phơi sáng hai lần, T không được lớn hơn nhiều so với độ dài hai sóng của các thành phần có tần số cao nhất trong bức ảnh. Điều này đảm bảo rằng sẽ không có nhiều thành phần trong khu vực chuyển tiếp. Trong khi nó có khả năng là các giới hạn này có thể vượt quá một chút mà không có sự thay đổi nào đáng chú ý. Vậy độ rộng dải tần của hình ảnh được đan nên xấp xỉ một quãng 8[5].

Làm thế nào để đan các ảnh có dải tần vượt quá một quãng 8? Cách tiếp cận được đề xuất ở đây là đầu tiên các ảnh này có thể được tách ra thành một tập các ảnh thành phần band-pass. Sau đó, mỗi phần đan với một T thích hợp. Cuối cùng, các thành phần band-pass đã đan được nối lại với nhau thành ảnh khảm mong muốn. Chúng ta

gọi hướng tiếp cận này là đan đa phân giải [5]. Kỹ thuật này đã được sử dụng để tạo nên hình 2.3f.

Trong việc phân tích ảnh thành nhiều dải tần số, điều quan trọng là dải tần trong ảnh gốc phải được phủ đều, mặc dù các dải tần bản thân chúng có thể chồng chéo lên nhau. Thực tế, tập các bộ lọc low-pass[5] - bộ lọc cho phép những ảnh có tần số bé hơn một tần số xác định đi qua được dùng để sinh một chuỗi các ảnh. Trong các ảnh này, giới hạn dải tần của chúng giảm dần qua từng ảnh với bước nhảy là một quãng 8. Các ảnh band-pass có thể được tạo ra dễ dàng bằng cách trừ mỗi ảnh low-pass cho ảnh ngay trước nó. Điều này không chỉ chắc chắn phủ hoàn toàn dải tần, mà còn có ý nghĩa là ảnh khảm cuối cùng có thể dễ dàng tạo ra bằng cách tổng hợp các ảnh band-pass thành phần.

Chuỗi các ảnh lọc low-pass G_0, G_1, \dots, G_N có thể được tạo ra bằng cách cuộn liên tiếp một hàm trọng số nhỏ với ảnh. Với kỹ thuật này, mật độ ảnh mẫu sẽ giảm đi sau mỗi lần lặp, do đó độ rộng dải tần cũng giảm trong các bước thống nhất một quãng 8. Sự giảm bớt mẫu cũng có nghĩa là chi phí tính toán sẽ giảm xuống mức tối thiểu.



Hình 2.4: Mô tả phương pháp lọc 1 chiều.

Dấu chấm của mỗi dòng trong hình 2.4 biểu diễn các mẫu, các điểm ảnh của một trong các ảnh đã được lọc. Dòng thấp nhất G_0 là ảnh gốc. Giá trị của mỗi node ở dòng tiếp theo (G_1) được tính bằng trung bình trọng số của 1 mảng con kích thước 5×5 của các node G_0 . Sau đó, các node của dòng G_2 được tính từ G_1 cũng với trọng số tương tự. Tiến trình được lặp đi lặp lại để tính G_2 từ G_1 , G_3 từ G_2 và cứ tiếp tục như vậy. Khoảng cách mẫu được nhân đôi sau mỗi lần lặp, do đó các mảng liên tiếp chỉ rộng bằng một nửa mảng trước nó. Nếu chúng ta hình dung các mảng này được xếp

mảng nọ chồng lên mảng kia thì kết quả sẽ là một cấu trúc dữ liệu hình chóp. Nếu kích thước ảnh gốc là $(2^N + 1) \times (2^N + 1)$ thì hình chóp sẽ có $N+1$ mức.

Cả mật độ mẫu và độ phân giải đều giảm theo từng mức của hình chóp. Trong trường hợp này, chúng ta sẽ gọi tiến trình sinh ra mỗi mức của hình chóp từ mức liền trước nó là phép toán REDUCE[5]. Đặt G_0 là ảnh gốc. Sau đó cho $0 < l < N$:

$$G_l = \text{REDUCE} [G_{l-1}] \quad (2.2)$$

$$\text{Nghĩa là: } G_l(i,j) = \sum \sum_{m,n=1}^5 w(m,n) G_{l-1}(2i+m, 2j+n) \quad (2.3)$$

Mẫu trọng số $w(m, n)$ được sử dụng để sinh mỗi mức hình chóp từ mức liền trước nó được gọi là nhân hàm sinh. Các trọng số này được chọn theo bốn điều kiện:

Đầu tiên, để thuận tiện cho tính toán, nhân hàm sinh phải tách được:

$$w(m, n) = \hat{w}(m) \hat{w}(n) \quad (2.4)$$

Thứ hai, mỗi hàm một chiều \hat{w} là đối xứng:

$$\hat{w}(0) = a, \hat{w}(-1) = \hat{w}(1) = b, \text{ và } \hat{w}(-2) = \hat{w}(2) = c \quad (2.5)$$

Thứ ba, \hat{w} thì được định mức: $a + 2b + 2c = 1$. (2.6)

Cuối cùng, mỗi node của mức l phải đóng góp cùng một trọng số tổng vào các node của mức $l+1$, vì thế $a + 2c = 2b$. Bây giờ, kết hợp các ràng buộc, chúng ta nhận thấy rằng a có thể coi là biến tự do, $b = 1/4$ và $c = 1/4 - a/2$.

2.1.1 Hàm trọng số tương đương

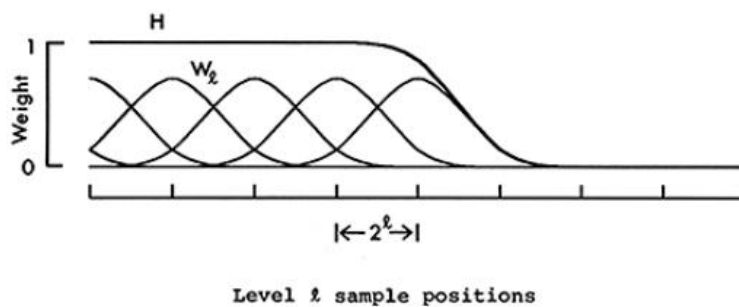
Rõ ràng là mỗi node mức l trong hình chóp tương ứng với một trung bình trọng số của mảng con 5×5 các node của mức $l-1$. Mỗi một lần lặp tương ứng với một giá trị trung bình của mảng con của mức $l-2$. Theo cách này, chúng ta có thể truy vết được các trọng số từ node đã cho của hình chóp ngược trở lại ảnh G_0 để nhận được “hàm trọng số tương đương”[5] W_l thoả mãn. Nếu được cuộn trực tiếp với ảnh gốc, nó sẽ cho các giá trị node giống mức l . Đó là thuận lợi để thảo luận về tính toán dựa trên hình chóp dưới dạng các hàm trọng số tương đương, mặc dù tiến trình lặp REDUCE hiệu quả hơn nhiều và được sử dụng trong tất cả các tính toán.

Các hàm trọng số tương đương có nhiều tính chất quan trọng trong các toán tử lọc và đan ghép. Tính co giãn của các hàm này tăng gấp đôi sau mỗi mức của hình chóp trong khi khuôn dạng của chúng không thay đổi. Khuôn dạng hàm phụ thuộc vào giá trị của các tham biến a trong nhân hàm sinh. Ví dụ, nếu $a = 0,5$ thì tất cả các hàm sẽ có dạng tam giác (triangular in shape). Nếu $a = 0,4$ thì các hàm sẽ tương tự với hàm mật độ xác suất Gauss. Việc cuộn với hàm Gauss sẽ cho hiệu ứng ảnh lọc low-pass. Cấu trúc hình chóp tương đương với cuộn ảnh với một tập các hàm giống như hàm Gauss để tạo ra tập các ảnh lọc tương ứng. Chúng ta sẽ xét chuỗi các ảnh G_0, \dots, G_N này như là hình chóp Gauss.

Giả sử các mẫu trong G_0 cách nhau một đơn vị khoảng cách. Sau đó, các mẫu tại mức l cách nhau khoảng 2^l . Điều này có thể được thể hiện bằng độ rộng của hàm trọng số tương đương W_l là $2^{l+2} - 4$, phủ $2^{l+2} - 3$ mẫu ảnh hoặc khoảng cách mẫu giảm 4 lần. Với mọi i, j và l :

$$\sum_{m,n=-2}^2 W_l(i - m2^l, j - n2^l) = 1 \quad (2.7)$$

Kết quả này có được từ tính chất phân bố đều của nhân hàm sinh. Khuôn dạng Gauss và các đặc tính tổng hợp của các hàm W_l có ý nghĩa là chúng ta có thể xây dựng hàm trọng số H cần tìm để đan ảnh (Hình 2.2). Giả sử rằng W_l liên quan tới mỗi node trong nửa trái của G_l trong khi trọng số bằng 0 với các node bên phải (hình 2.5). Sau đó, tổng của các hàm này sẽ tạo thành một hàm điều giảm từ 1 xuống 0 với miền chuyển tiếp độ rộng T gấp 3 lần khoảng cách mẫu mức l . Tính chất này sẽ được sử dụng trong kĩ thuật đan đa phân giải dựa trên hình chóp (pyramid-based multiresolution spline), mặc dù các hàm H và W không bao giờ tính được một cách cụ thể.



Hình 2.5: Hàm trọng số tương đương

2.1.2 Hình chóp Laplace

Hình chóp Gaussian là tập các ảnh lọc low-pass. Để tạo ra các ảnh band-pass cần thiết cho đa phân giải (multiresolution spline), chúng ta trừ mỗi mức của hình chóp cho mức thấp nhất trước nó. Vì các mảng này khác biệt về mật độ mẫu, nên trước khi trừ cần phải nội suy các mẫu trung gian. Phép nội suy có thể đạt được bằng cách đảo chiều tiến trình REDUCE. Chúng ta sẽ gọi phép toán tử này là EXPAND[5]. Đặt $G_{l,k}$ là ảnh thu được bằng cách mở rộng G_l k lần:

$$G_{l,0} = G_l \quad (2.8)$$

Và với $k > 0$:

$$G_{l,k} = \text{EXPAND}[G_{l,k-1}] \quad (2.9)$$

Với EXPAND, chúng ta đã thực hiện:

$$G_{l,k}(i, j) = 4 \sum_{m,n=-2}^2 G_{l,k-1}\left(\frac{2i+m}{2}, \frac{2j+n}{2}\right) \quad (2.10)$$

Ở đây, $(2i+m)/2$ và $(2j+n)/2$ là các số nguyên và $G_{l,l}$ cùng kích thước với G_{l-1} và $G_{l,l}$ cùng kích thước với ảnh gốc. Bây giờ, chúng ta sẽ định nghĩa một chuỗi ảnh band-pass L_0, L_1, \dots, L_N . với $0 < l < N$

$$L_l = G_l - \text{EXPAND}[G_{l+1}] = G_l - G_{l+1,1} \quad (2.11)$$

Vì không có mảng mức cao hơn để trừ cho G_N , nên chúng ta định nghĩa $L_N = G_N$. Khi giá trị của mỗi node trong hình chóp Gaussian thu được trực tiếp từ việc cuộn hàm trọng số W_l với ảnh. Mỗi node của L_l có thể thu được trực tiếp bằng cách cuộn $W_l - W_{l+1}$ với ảnh. Sự khác biệt của hàm Gaussian-like[5] giống như những người thường sử dụng phương pháp Laplace trong xử lý ảnh. Vì vậy chúng ta có thể tham khảo các dãy L_0, L_1, \dots, L_N như là phương pháp hình chóp laplace.

Các bước tạo ra hình chóp Laplace cũng có thể được đảo ngược để khôi phục lại ảnh gốc G_0 một cách chính xác. Mức trên cùng của chóp, L_N , trước tiên được mở rộng và cộng vào L_{N-1} để khôi phục G_{N-1} . Mảng này sau đó được mở rộng và cộng vào L_{N-2} để được G_{N-2} và cứ tiếp tục như vậy. Ta có thể viết:

$$G_0 = \sum_{l=0}^N L_{l,l} \quad (2.12)$$

Các biện pháp mở rộng và tổng hợp có thể được sử dụng để tạo ảnh khảm từ tập các thành phần band-pass của chính nó.

2.1.3 Các điều kiện đường biên

Cả hai phép toán REDUCE và EXPAND cần đặc biệt chú ý tới các nút biên. Ví dụ, khi một thủ tục REDUCE được thực thi, nhân hàm sinh cho mỗi node biên tại mức G_{l+1} mở rộng thêm hai node so với G_l . Vì thế, trước khi thực thi REDUCE hay EXPAND, G_l được tăng lên hai hàng node mỗi chiều. Các giá trị được gán tại các nút cho bởi phép phản xạ và đảo ngược qua các node biên. Vì thế, nếu $G_l(0, j)$ là node biên trái của G_l , chúng ta đặt:

$$G_l(-1, j) = 2G_l(0, j) - G_l(1, j) \quad (2.13)$$

Và:
$$G_l(-2, j) = 2G_l(0, j) - G_l(2, j) \quad (2.14)$$

2.1.4 Kỹ thuật đan đa phân giải

2.1.4.1 Đan chồng các ảnh

Thuật toán đan đa phân giải có thể được định nghĩa đơn giản hơn dưới dạng các toán tử hình chóp cơ bản đã được giới thiệu ở trên. Thuật toán có thể thay đổi để đan các ảnh có một phần chung (gối lên nhau - overlap), không có phần chung (không gối lên nhau - nonoverlap) và đan các ảnh có hình dạng tùy ý. Để bắt đầu, giả sử chúng ta muốn đan nửa trái ảnh A với nửa phải ảnh B . Giả sử các ảnh này đều là hình vuông, mỗi cạnh $2^N + 1$ pixel mỗi chiều và giả sử chúng hoàn toàn chồng lấp. Quá trình đan thực hiện sau 3 bước:

Bước 1: Dựng các hình chóp Laplace LA và LB tương ứng cho ảnh A và ảnh B .

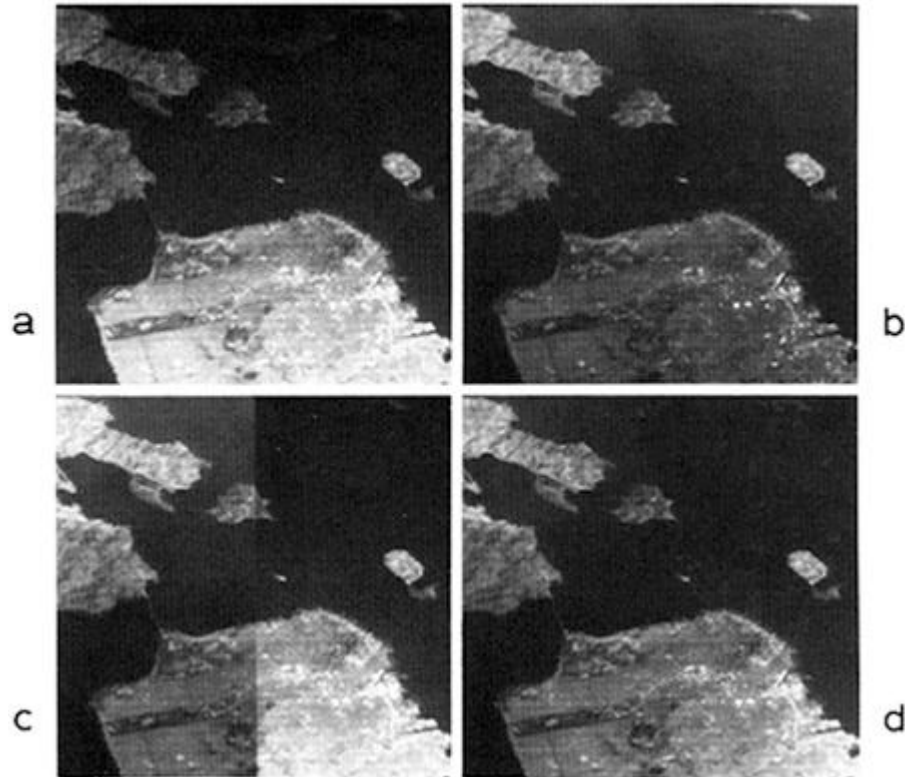
Bước 2: Dựng một chóp Laplace LS bằng cách sao chép các cạnh từ bên trái của LA vào các node tương ứng của LS và các node bên phải LB vào bên phải LS . Các node thuộc đường chính giữa của LS được gán bằng giá trị trung bình của các node LA và LB tương ứng.

Đường chính giữa tại mức l của chóp Laplace tại $i=2^{N-1}$. Vì vậy, với mọi i, j, l :

$$LS_l(i, j) = \begin{cases} LA_l(i, j) & \text{if } i < 2^{N-1} \\ LA_l(i, j) + LB_l(i, j) & \text{if } i = 2^{N-1} \\ LB_l(i, j) & \text{if } i > 2^{N-1} \end{cases} \quad (2.15)$$

Bước 3: Ảnh đan S được tạo bằng cách mở rộng và cộng các mức của LS .

Kết quả của việc áp dụng thủ tục này cho ảnh ngôi sao trên được trình bày trong hình 2.3f.



Hình 2.6: Ảnh ghép từ hai ảnh vệ tinh của San Francisco.

Ví dụ thứ 2 được thể hiện trong hình 2.6. Ở đây, chúng ta muốn đan hai ảnh vệ tinh của San Francisco, hình 2.6a và hình 2.6b. Các ảnh này giống hệt nhau trừ độ nhiễu nền khuyếch tán do khác nhau do điều kiện khí quyển hoặc quá trình xử lý ảnh. Một lần nữa, chúng ta muốn tạo nên một bức ảnh ghép mà trong đó, nửa bên trái là của một ảnh và nửa bên phải là của ảnh khác. Nếu chỉ ghép mà không đan ảnh thì đường biên sẽ lộ rõ như hình 2.6c. Tuy nhiên, nếu sử dụng kỹ thuật đan đa phân giải, cạnh hầu như bị xoá bỏ, điều này được thể hiện ở hình 2.6d.

2.1.4.2 Đan các ảnh với các miền tùy ý

Các bước ở trên có thể được tổng quát hoá lên để tạo ra một ảnh khảm từ các miền ảnh có hình dạng tùy ý. Một lần nữa, chúng ta giả thiết rằng miền được đan chứa trong ảnh A và B . Như trên, các node của các chóp Laplace LA và LB cho các ảnh thành phần sẽ được kết nối để tạo thành chóp Laplace LS của ảnh kết quả S . Chúng ta giới thiệu thêm một cấu trúc tháp để xác định các node nào của LS được lấy từ LA , node nào từ LB và node nào là trung bình của hai. Đặt R là ảnh nhị phân kích thước bằng A và B , trong đó, tất cả các điểm ảnh trong miền của ảnh A được đan với B có giá trị là 1 và tất cả các điểm ngoài miền này là 0. Các bước của kỹ thuật đan đa phân giải được chỉnh sửa như sau:

Bước 1:

- Dựng các chóp Laplace LA và LB tương ứng cho ảnh A và B .
- Dựng chóp Gaussian GR cho miền ảnh R

Bước 2:

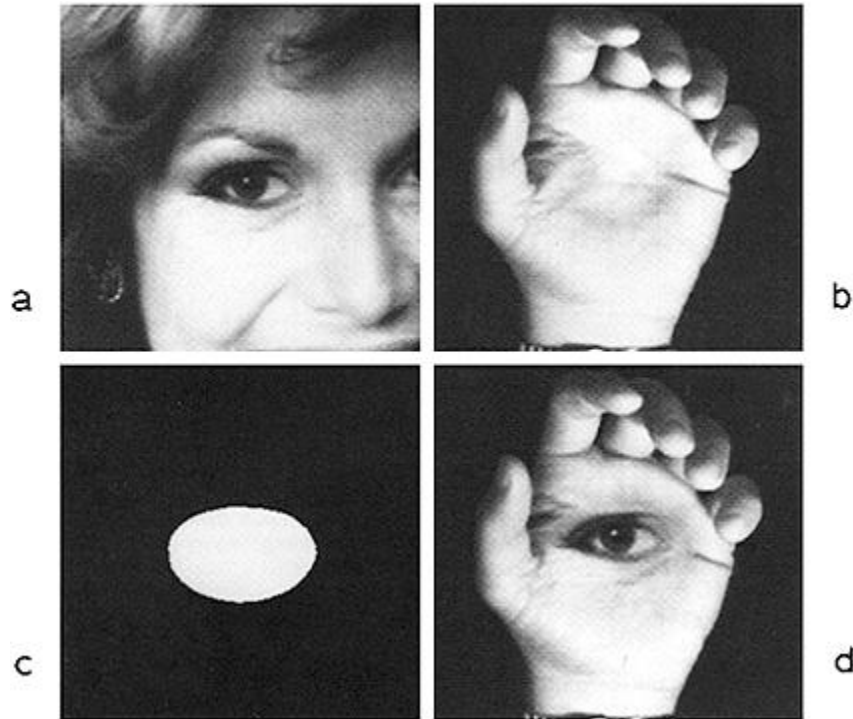
- Tạo một chóp liên kết LS từ LA và LB sử dụng các node của GR như là các trọng số. Nghĩa là, với mỗi l, i, j :

$$LS_l(i, j) = GR_l(i, j)LA_l(i, j) + (1 - GR_l(i, j))LB_l(i, j) \quad (2.16)$$

Bước 3:

- Thu được ảnh đan S bằng cách mở rộng và cộng các mức của LS .

Hình chóp Gaussian phục vụ hai mục đích: nó là một phương thức tiện lợi để xác định các node tại mỗi mức nằm trong vùng mặt nạ của ảnh R và nó “làm mềm” các cạnh của mặt nạ trong suốt quá trình lọc. Các node rơi đúng vào biên mặt nạ sẽ nhận trọng số 50%.



Hình 2.7: Đan hình ảnh con mắt vào trong lòng bàn tay.

Trong hình 2.7, hình 2.7a và hình 2.7b là 2 ảnh gốc: một con mắt và một bàn tay. Hình 2.7c là miền của bức ảnh đầu tiên được đan vào bức ảnh thứ hai. Hình 2.7d là kết quả cuối cùng của kỹ thuật đan: Con mắt đã được đan vào trong lòng bàn tay

2.1.4.3 Đan các ảnh không chồng lên nhau

Kỹ thuật trung bình trọng số chỉ được áp dụng cho các bức ảnh chồng lên nhau. Tuy nhiên, một ảnh đan phù hợp có thể được tạo bởi các bức ảnh tiếp giáp với nhau nhưng không chồng lên nhau nếu mỗi ảnh được ngoại suy qua biên của nó để tạo thành một miền chuyển tiếp chồng lên nhau. Vì độ rộng của miền chuyển tiếp có thể là một phần của độ rộng của bản thân bức ảnh, nên việc ngoại suy dường như là một nhiệm vụ khó khăn. Tuy nhiên, trong kỹ thuật đan đa phân giải ngoại suy có thể thực hiện độc lập trong mỗi dải tần. Hơn nữa, khi thuật toán hình chóp được sử dụng, chỉ có hai ảnh mẫu cần được ngoại suy bên ngoài biên của mỗi mức. Trong thực tế nó chỉ là loại ngoại suy của đường biên đã được xử lý trong khi xây dựng cả hình chóp Gaussian và Laplace. Không có các bước tiếp theo cần được thực hiện trong đan ghép.



Hình 2.8: Ứng dụng của việc đan các ảnh không chồng nhau

Chúng ta bắt đầu với một ảnh duy nhất mà chính nó là một ảnh khảm của khối 16×16 pixel (chỉ có mảng trung tâm 8×8 của khối được biểu diễn trong hình minh họa). Mỗi khối được xây dựng lại từ một sự biến đổi mã rất cao. Trong trường hợp này, đại diện cho hình ảnh ở tỉ lệ chỉ 0.5 bit/pixel . Khối chuyển đổi mã ở mức rất thấp của khối biên nổi bật tạo ra trong phép cộng để suy biến mạnh các hình ảnh khác. Nhiệm vụ của chúng ta là loại bỏ đường biên bằng phương pháp đan đa phân giải.

Ảnh gốc có kích thước 16×16 của khối. Trước khi thử nghiệm đan, chúng ta sử dụng phép ngoại suy để thêm một hàng bên phải và bên dưới của mỗi khối. Kết quả các khối 17×17 phù hợp với cấu trúc hình chóp và sự ghép chồng vào ảnh bởi 1 pixel mỗi chiều. Đây là bước đầu tiên của đan ghép. Phương pháp là xây dựng một hình chóp Laplace cho mỗi khối của 256 khối. Những hình chóp này được nối lại thành một hình chóp duy nhất với với các nút trong cạnh chồng lên nhau và được lấy trung bình của các mức.

Đặt L_{lmn} là mức l th của hình chóp laplace đã được xây dựng, với khối n th trong hàng thứ m th của khối. Sau đó, với i và j không trong khối đường biên (ví dụ: i và j không bằng bội số của 2^{4-l}) và $0 < l < 3$

$$LS_l(i, j) = L_{lmn}(\hat{i}, \hat{j}) \quad (2.17)$$

Trong đó m là phần nguyên của $i/2^{4-l}$, n là phần nguyên của

$$j/2^{4-l}, \hat{i} = i - m2^{4-l}, \text{ và } \hat{j} = j - n2^{4-l} \quad (2.18)$$

Nếu i hoặc j là một khối ranh giới trên mức 0 , $0 < l < 3$, sau đó nút LS sẽ được lấy trung bình.

$$LS_l(i, j) = \frac{L_{l,m-1,n}(16, \hat{j}) + L_{l,m,n}(0, \hat{j})}{2} \quad (2.19)$$

Giá trị trung bình này không được tính cho các nút đường biên ở mức dưới cùng bởi nút $L_{l,m-1,n}(16, j)$ đại diện cho một giá trị ngoại suy. Thay vào đó, chúng ta chỉ nói đơn giản:

$$LS_0(i, j) = L_{0,m,n}(0, \hat{j}) \quad (2.20)$$

Khôi phục lại thông qua quá trình mở rộng và tổng hợp được thấy trong hình 2.8b. Lưu ý rằng các khối đường biên gần như đã được loại bỏ hoàn toàn. Hình ảnh vẫn ở mức chất lượng thấp, nhưng điều này là do tỉ lệ bit ban khối ban đầu trước khi mã hóa so với kĩ thuật đan ghép.

2.2 TRÍCH CHỌN ĐẶC TRƯNG TRONG KHẢM ẢNH

2.2.1 Kĩ thuật trích chọn đặc trưng

Một bức ảnh mang rất nhiều thông tin như định dạng, kích thước, màu sắc, độ tương phản v.v... Trong khảm ảnh với một tập ảnh mẫu rất lớn thì việc tra cứu các ảnh mẫu này với các vùng trên ảnh gốc gặp rất nhiều khó khăn. Vì vậy việc áp dụng kĩ thuật trích chọn đặc trưng sẽ giúp ích rất nhiều cho việc khảm ảnh. Thông thường đối với việc khảm ảnh thì chúng ta sẽ quan tâm nhiều đến một số đặc trưng mà Tee Cheng Siew[4]] và Lowe David [3] đã chỉ ra một số đặc trưng như sau:

Đặc trưng màu sắc: Màu sắc là một đặc trưng nổi bật và được sử dụng phổ biến nhất trong tìm kiếm ảnh theo nội dung. Mỗi một điểm ảnh (thông tin màu sắc) có thể được biểu diễn như một điểm trong không gian màu sắc ba chiều. Các không gian màu sắc thường dùng là: RGB, CIE, HSV... Tìm kiếm ảnh theo màu sắc tiến hành tính toán biểu đồ màu cho mỗi ảnh để xác định tỉ trọng các điểm ảnh của ảnh mà chứa các giá trị màu. Các nghiên cứu gần đây đang cố gắng phân vùng ảnh theo các màu sắc khác nhau và tìm mối quan hệ giữa các vùng này.

Đặc trưng kết cấu: Trích xuất nội dung ảnh theo kết cấu nhằm tìm ra mô hình trực quan của ảnh và cách thức chúng được xác định trong không gian. Kết cấu được biểu diễn bởi các texel mà sau đó được đặt vào một số các tập phụ thuộc vào số kết cấu

được phát hiện trong ảnh. Các tập này không chỉ xác định các kết cấu mà còn chỉ rõ vị trí các kết cấu trong ảnh. Việc xác định các kết cấu đặc biệt trong ảnh đạt được chủ yếu bằng cách mô hình các kết cấu như những biến thể cấp độ xám 2 chiều.

Đặc trưng hình dạng: Hình dạng của một ảnh hay một vùng là một đặc trưng quan trọng trong việc xác định và phân biệt ảnh trong nhận dạng mẫu. Mục tiêu chính của biểu diễn hình dạng trong nhận dạng mẫu là đo thuộc tính hình học của một đối tượng được dùng trong phân lớp, so sánh và nhận dạng đối tượng.

2.2.1.1 Đặc trưng màu sắc

Tra cứu ảnh theo lược đồ màu là phương pháp phổ biến và được sử dụng nhiều nhất trong các hệ thống tìm kiếm ảnh theo nội dung. Đây là phương pháp đơn giản, tốc độ tìm kiếm tương đối nhanh tuy nhiên kết quả tìm kiếm có độ chính xác không cao. Đây có thể xem là bước lọc đầu tiên cho những bước tìm kiếm sau. Một số lược đồ màu được sử dụng như: lược đồ màu RGB, lược đồ màu HSI, lược đồ HSI cải tiến. Trong đó, lược đồ màu RGB được sử dụng phổ biến nhất

- **Hệ màu RGB:**

Đối với ảnh 256 màu, lược đồ màu của ảnh tương đương với lược đồ màu của ảnh xám. Đối với ảnh 24 bit màu, lược đồ miêu tả khả năng kết nối về cường độ của ba kênh màu R, G, B. Lược đồ màu này được định nghĩa như sau:

$$h_{R,G,B}[r,g,b] = N * \text{Prob}\{R=r,G=g,B=b\} \quad (2.21)$$

Trong đó N là số lượng điểm ảnh. Lược đồ màu này được tính bằng cách rời rạc hóa từng màu trong ảnh, sau đó đếm số điểm ảnh của mỗi màu. Khi mà số lượng màu là có hạn, để thuận tiện hơn, người ta thường chuyển đổi ba kênh màu thành một biến giá trị duy nhất. Một cách khác để tính lược đồ màu của ảnh RGB là ta phân ra làm 3 lược đồ riêng biệt $h_R[]$, $h_G[]$, $h_B[]$. Khi đó, mỗi lược đồ được tính bằng cách đếm kênh màu tương ứng trong mỗi điểm ảnh.

- **Độ đo tương đồng về màu sắc**

Một số độ đo tương đồng được sử dụng như: Độ đo khoảng cách Öclit, độ đo Jensen-Shannon divergence (JSD). Gọi $h(I)$ và $h(M)$ tương ứng là 2 lược đồ màu của hai ảnh I và ảnh M. Khi đó các loại độ đo màu được định nghĩa là một số nguyên (hoặc số thực) theo các loại độ đo tương ứng như sau:

❖ Khoảng cách Euclid:

$$\text{Intersection}(h(I), h(M)) = \sum_{j=1}^K \sqrt{(h(I) - h(M))^2} \quad (2.22)$$

$$\text{Hoặc: } \text{Intersection}(h(I), h(M)) = \sum_{j=1}^K |h(I) - h(M)| \quad (2.23)$$

❖ Độ đo Jensen-Shannon divergence:

Độ đo Jensen-Shannon divergence sử dụng lược đồ màu RGB để tính toán độ tương đồng về màu sắc giữa 2 ảnh:

$$d_{\text{JSD}}(H, H') = \sum_{m=1}^M H_m \log \frac{2H_m}{H_m + H'_m} + H'_m \log \frac{2H'_m}{H_m + H'_m} \quad (2.24)$$

Trong đó: H, H' là 2 biểu đồ màu được so sánh, H_m là bin thứ m của biểu đồ H

❖ Độ đo khoảng cách Min-Max

Được thực hiện dựa trên ý tưởng lấy phần giao của 2 lược đồ màu cần so sánh, ta sẽ được một lược đồ, tính tổng các giá trị có được từ lược đồ này cho ta được độ đo min-max.

- Đối với độ đo min ta tính dựa vào giá trị min tại khác K bin:

$$\text{Intersection}(h(I), h(M)) = \sum_{j=1}^K \min\{H(I)[j], H(M)[j]\} \quad (2.25)$$

- Tương tự đối với độ đo Max:

$$\text{Intersection}(h(I), h(M)) = \sum_{j=1}^K \text{Max}\{H(I)[j], H(M)[j]\} \quad (2.26)$$

$$\text{Matching}(h(I), h(M)) = \frac{\text{Intersection}(h(I), h(M))}{\text{Max}(\sum h(I)[j], \sum h(M)[j])} \quad (2.27)$$

2.2.1.2 Đặc trưng kết cấu

Kết cấu cung cấp thông tin về sự sắp xếp về mặt không gian của màu sắc và cường độ một ảnh. Kết cấu được đặc trưng bởi sự phân bố không gian của những mức cường độ trong một khu vực lân cận với nhau. Kết cấu gồm các kết cấu gốc hay nhiều kết cấu gộp lại đôi khi gọi là texel. Có một số phương pháp dụng để trích xuất các đặc trưng kết cấu như :

- Phương pháp hình chóp
- Biến đổi đường viền
- Biến đổi sóng Gabor

- Biểu diễn ma trận đồng hiện

2.2.1.3 Đặc trưng hình dạng

Màu sắc và kết cấu là những thuộc tính mang tính toàn cục trong một ảnh. Trong khi đó, hình dạng không phải là một thuộc tính của ảnh. Nói tới hình dạng của hình ảnh là chỉ đến một khu vực đặc biệt trong ảnh, hay hình dạng chỉ là biên của một đối tượng nào đó trong ảnh.

Trong tra cứu ảnh theo nội dung, hình dạng là một cấp cao hơn so với màu sắc và kết cấu. Nó đòi hỏi sự phân biệt giữa các vùng để tiến hành xử lý về độ đo của hình dạng. Các hệ thống tra cứu ảnh theo nội dung thường khai thác hai nhóm biểu diễn hình dạng sau :

- Biểu diễn hình dạng theo đường biên (cotour-based descriptor): Biểu diễn các đường biên bao bên ngoài.
- Biểu diễn theo vùng (region-based descriptor): Biểu diễn một vùng toàn vẹn.

Độ đo về hình dạng rất nhiều trong phạm vi lý thuyết của bộ môn xử lý ảnh. Chúng trải rộng từ những độ đo toàn cục dạng thô với sự trợ giúp của việc nhận dạng đối tượng, cho tới những độ đo chi tiết tự động tìm kiếm những hình dạng đặc biệt. Lược đồ hình dạng là một ví dụ của độ đo đơn giản. Kỹ thuật dùng đường biên hiệu quả hơn phương pháp trước, chúng tìm kiếm những hình dạng đối tượng gần giống với đường biên nhất. Phương pháp vẽ phác họa là phương pháp có nhiều đặc trưng rõ ràng hơn, không chỉ tìm kiếm những đường biên đối tượng đơn, mà còn đối với tập những đối tượng đã được phân đoạn trong một ảnh mà người dùng vẽ hay cung cấp.

2.2.1.4 Đặc trưng cục bộ bất biến

Người ta thường chia đặc trưng cục bộ thành 2 loại là những điểm trích xuất được từ “điểm nhô ra” (salient points) của ảnh và đặc trưng SIFT được trích chọn từ các điểm hấp dẫn Haris (interest points). Trong phần này, chúng ta sẽ trình bày chi tiết về việc trích chọn các đặc trưng cục bộ bất biến (Scale Invariant Feature Transform) của ảnh[3].

Các đặc trưng này bất biến với việc thay đổi tỉ lệ ảnh, quay ảnh, đôi khi là thay đổi điểm nhìn và thêm nhiễu ảnh hay thay đổi cường độ chiếu sáng của ảnh. Phương pháp được lựa chọn có tên là Scale-Invariant Feature Transform và đặc trưng trích rút

được gọi là đặc trưng SIFT (SIFT Feature). Các đặc trưng SIFT này được trích rút ra từ các điểm hấp dẫn cục bộ (Local Interest Point)

Phương pháp trích chọn các đặc trưng bất biến SIFT được tiếp cận theo phương pháp thác lọc, theo đó phương pháp được thực hiện lần lượt theo các bước sau[3]:

- **Phát hiện các điểm cực trị Scale-Space** (Scale-Space extrema detection): Bước đầu tiên này tiến hành tìm kiếm các điểm hấp dẫn trên tất cả các tỉ lệ và vị trí của ảnh. Nó sử dụng hàm different-of-Gaussian để xác định tất cả các điểm hấp dẫn tiềm năng mà bất biến với quy mô và hướng của ảnh.
- **Định vị các điểm hấp dẫn** (keypoint localization): Một hàm kiểm tra sẽ được đưa ra để quyết định xem các điểm hấp dẫn tiềm năng có được lựa chọn hay không?
- **Xác định hướng cho các điểm hấp dẫn** (Orientation assignment): Xác định hướng cho các điểm hấp dẫn được chọn
- **Mô tả các điểm hấp dẫn** (Keypoint descriptor): Các điểm hấp dẫn sau khi được xác định hướng sẽ được mô tả dưới dạng các vector đặc trưng nhiều chiều.

❖ Phát hiện điểm cực trị Scale-space

Các điểm hấp dẫn với đặc trưng SIFT tương thích với các cực trị địa phương của bộ lọc difference-of-Gaussian (DoG) ở các tỉ lệ khác nhau. Định nghĩa không gian tỉ lệ của một hình ảnh là hàm $L(x,y,k\sigma)$ được mô tả như sau:

$$L(x,y,\sigma) = G(x,y,k\sigma) * I(x,y) \quad (2.28)$$

Với : $G(x,y,k\sigma)$: biến tỉ lệ Gaussian (variable scale Gaussian)

$I(x,y)$: Ảnh đầu vào

* là phép nhân chập giữa x và y

$$\text{Và : } G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2.29)$$

Để phát hiện được các điểm hấp dẫn, ta đi tìm các cực trị của hàm DoG được định nghĩa:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) \quad (2.30)$$

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (2.31)$$

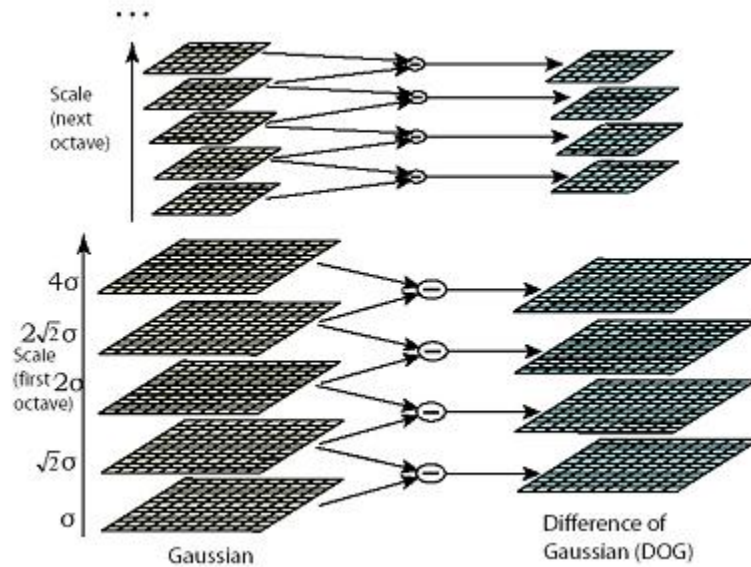
Giá trị hàm DoG được tính xấp xỉ dựa vào giá trị Scale-normalized Laplacian of Gaussian ($\sigma^2 \nabla^2 G$) thông qua các phương trình (2.28),(2.29),(2.31)

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G \quad (2.32)$$

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x,y,k\sigma) - G(x,y,\sigma)}{k} \quad (2.33)$$

$$G(x,y,k\sigma) - G(x,y,\sigma) \approx (k - 1)\sigma \nabla^2 G \quad (2.34)$$

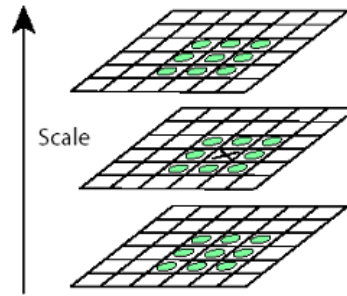
Như vậy, bước đầu tiên của giải thuật SIFT phát hiện các điểm hấp dẫn với bộ lọc Gaussian ở các tỉ lệ khác nhau và các ảnh GoG từ sự khác nhau của các ảnh mờ.



Hình 2.9: Biểu đồ mô phỏng việc tính toán các DoG ảnh từ các ảnh mờ

Các ảnh cuộn được nhóm thành các octave (mỗi octave tương ứng với giá trị gấp đôi của σ). Giá trị của k được chọn sao cho số lượng ảnh đã làm mờ (blurred images) cho mỗi octave là cố định. Điều này đảm bảo cho số lượng các ảnh DoG cho mỗi octave không thay đổi.

Các điểm hấp dẫn được xác định là các cực đại hoặc cực tiểu của các ảnh DoG qua các tỉ lệ. Mỗi điểm ảnh trong DoG được so sánh với 8 điểm ảnh láng giềng của nó ở cùng tỉ lệ đó và 9 láng giềng kề ở các tỉ lệ ngay trước và sau nó. Nếu điểm ảnh đó đạt giá trị cực tiểu hoặc cực đại thì sẽ được chọn làm các điểm hấp dẫn phù hợp

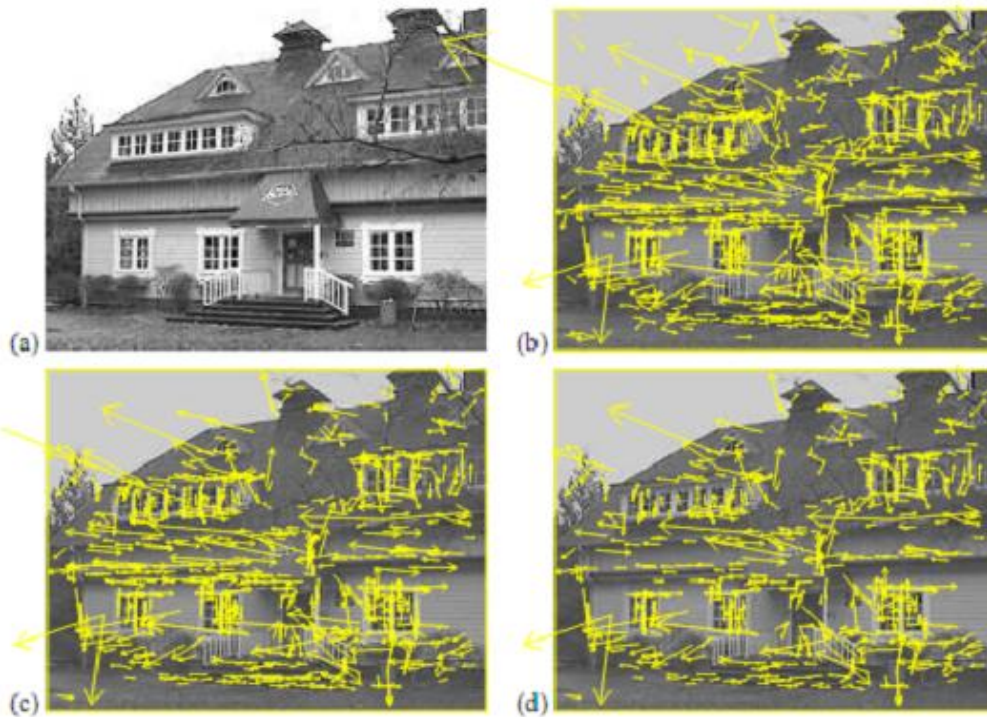


Hình 2.10: Mỗi điểm ảnh được so sánh với 26 láng giềng của nó.

❖ Định vị điểm hấp dẫn

Mỗi điểm hấp dẫn sau khi được chọn sẽ được đánh giá xem có được giữ lại hay không:

- Loại bỏ các điểm hấp dẫn có độ tương phản thấp
- Một số điểm hấp dẫn dọc theo các cạnh không giữ được tính ổn định khi ảnh bị nhiễu cũng bị loại bỏ. Các điểm hấp dẫn còn lại sẽ được xác định hướng.



Hình 2.11: Quá trình lựa chọn các điểm hấp dẫn

❖ Xác định hướng cho điểm hấp dẫn

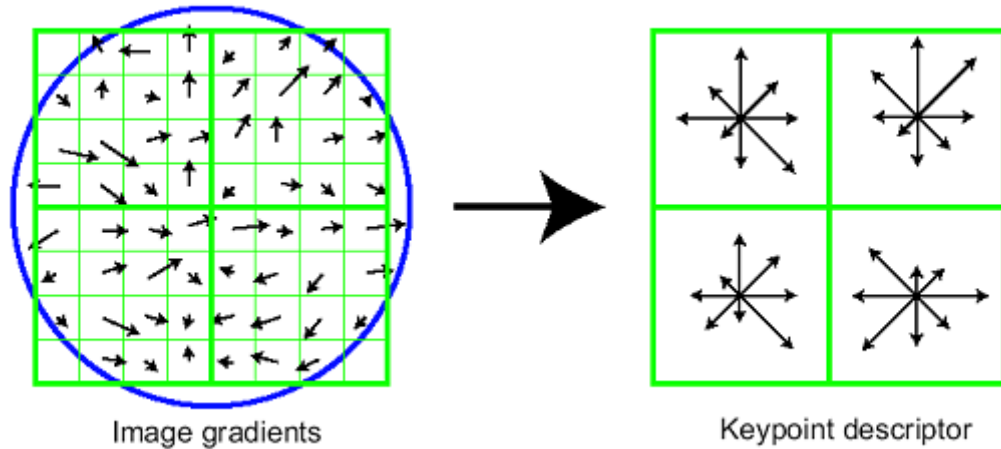
Để xác định hướng cho các điểm hấp dẫn, người ta tính toán biểu đồ hướng Gradient trong vùng láng giềng của điểm hấp dẫn. Độ lớn và hướng của các điểm hấp dẫn được xác định theo công thức:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (2.35)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))) \quad (2.36)$$

❖ Biểu diễn vector cho điểm hấp dẫn

Điểm hấp dẫn sau khi được xác định hướng sẽ được biểu diễn dưới dạng các vector $4 \times 4 \times 8 = 128$ chiều.



Hình 2.12: Biểu diễn các vector đặc trưng

❖ Đo độ tương đồng cho đặc trưng cục bộ bất biến

Một số độ đo tương đồng cho ảnh sử dụng đặc trưng SIFT như :

- Độ đo Cosin:

$$d(x,y) = \frac{x,y}{\|x\| \cdot \|y\|} \quad (2.37)$$

- Khoảng cách góc:

$$d(x,y) = \cos^{-1}(x,y) \quad (2.38)$$

- Độ đo Euclide:

$$d(x,y) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2} \quad (2.39)$$

- Độ đo Jensen-Shannon divergence:

$$d_{\text{JSD}}(H,H') = \sum_{m=1}^M H_m \log \frac{2H_m}{H_m + H'_m} + H'_m \log \frac{2H'_m}{H_m + H'_m} \quad (2.40)$$

Với H, H' là 2 biểu đồ biểu diễn các vector đặc trưng SIFT

Lựa chọn đặc trưng

Sau khi trích chọn được các đặc trưng nội dung của ảnh, tập các đặc trưng có thể được tối ưu hóa bằng các phương pháp lựa chọn đặc trưng để tăng chất lượng và hiệu quả khi sử dụng các tập đặc trưng.

Một cách tổng quát, lựa chọn đặc trưng là phương pháp giảm thiểu các đặc trưng nhằm chọn ra một tập con các đặc trưng phù hợp trong học máy để xây dựng mô hình học tốt nhất. Mục đích của lựa chọn đặc trưng là tìm ra không gian con các đặc trưng tối ưu sao cho các tập ảnh “thích hợp” và “không thích hợp” được tách biệt nhất.

Có nhiều phương pháp lựa chọn đặc trưng được đề xuất như: phương pháp tăng khuếch đại (boosting manner) kết hợp với nền tảng Real Adaboost của Wei Jian và Guihua Er. Mingjing Li đưa ra tiêu chí lựa chọn các đặc trưng là: Mô hình tương phản đặc trưng được tổng quát hóa (Generalized Feature Contrast Model) dựa trên mô hình tương phản đặc trưng (Feature Contrast Model). Một số phương pháp cổ điển khác như phương pháp dựa vào phân phối (distribution based). Phương pháp dựa vào phân tích biệt thức (Discriminant analysis DA) ví dụ như phân tích đa biệt thức (Multiple Discriminant analysis MDA), phân tích biệt thức không đối xứng (biased Discriminant analysis BDA). Phương pháp tối đa khoảng cách tối thiểu đối xứng trong không gian con (symmetric maximized minimal distance in subspace SMMS)...

Có nhiều phương pháp để đánh giá kết quả của tập con các đặc trưng. Vì vậy, kết quả đối với những mô hình lựa chọn đặc trưng khác nhau là khác nhau. Hai mô hình phổ biến cho lựa chọn đặc trưng là: Mô hình Filter và mô hình Wrapper[4].

- **Mô hình Filter:** đánh giá mỗi phần tử bằng một vài tiêu chuẩn hay độ đo nào đó, rồi chọn ra tập con các thuộc tính được đánh giá cao nhất.

- **Mô hình Wrapper:** Sử dụng một thuật toán tìm kiếm để đánh giá tập con các thuộc tính coi như là một nhóm hơn là một phần tử riêng lẻ. Cốt lõi của mô hình Wrapper là một thuật toán học máy cụ thể. Nó đánh giá độ tốt của những tập con đặc trưng tùy theo độ chính xác học của tập con, điều này xác định thông qua một tiêu chí nào đó.

2.2.2 Ứng dụng kỹ thuật trích chọn đặc trưng trong khám ảnh

Khác với kỹ thuật đơn đa phân giải là ứng dụng để khám ảnh toàn cảnh, kỹ thuật trích chọn đặc trưng ảnh được ứng dụng để khám ảnh nhiều lớp. Việc cần làm đối với ảnh khám nhiều lớp đó là sử dụng các lưới để chia ảnh nguồn ra nhiều phần bằng nhau và mỗi phần này chính là kích thước của các tập ảnh mẫu. Tuy nhiên không phải lúc nào cũng như vậy. Đối với một số kỹ thuật nâng cao thì các lưới này sẽ có dạng tùy ý dựa vào kết cấu của ảnh nguồn. Ví dụ như ảnh 1.21 ảnh quảng cáo của hãng IKA. Ảnh được tạo thành từ nhiều mẫu có kích thước khác nhau. Với những ảnh kiểu này đòi hỏi kỹ thuật rất cao. Hiện giờ đa số các phần mềm khám ảnh đều chia ảnh bởi các lưới đều nhau là các hình vuông hoặc chữ nhật.

Quy trình khám ảnh gồm các bước sau:

Bước 1: Chuẩn bị ảnh nguồn (nên chỉnh lại cỡ ảnh sao cho kích thước nó là bội số của kích thước các tập ảnh mẫu) và tập ảnh mẫu (tập ảnh mẫu này được lưu cùng một thư mục)

Bước 2: Chia ảnh nguồn bằng các lưới, các lưới này là các lưới có mắt hình vuông hoặc chữ nhật, và lưới này là bội số của các mẫu. Lưới có kích thước càng nhỏ thì ảnh khám càng có chất lượng tốt.

Bước 3: Duyệt từng vùng (các mắt lưới) của ảnh nguồn. Sử dụng kỹ thuật trích chọn đặc trưng để so sánh từng vùng đó với tất cả các ảnh trong tập ảnh mẫu để tìm ra ảnh mẫu phù hợp nhất (có màu sắc tương ứng, kết cấu tương ứng ...). Gán ảnh mẫu phù hợp vào vị trí của vùng tương ứng.

Vì khám ảnh nhiều lớp được ghép từ nhiều ảnh có nội dung khác nhau. Nên không giống như ảnh khám toàn cảnh, nó không cần phải làm mịn đường biên và cân bằng màu sắc, độ tương phản,... nên có thể nói ở bước này khám ảnh nhiều lớp nhàn hơn một chút.

CHƯƠNG 3: CHƯƠNG TRÌNH THỬ NGHIỆM

3.1 Bài toán

Chương trình khảm ảnh được xây dựng để minh họa cho phương pháp khảm ảnh nhiều lớp dựa trên việc trích chọn đặc trưng của ảnh số.

- Đầu vào : Một ảnh lớn và một tập ảnh nhỏ dùng làm mẫu.
- Đầu ra: Ảnh lớn được ghép từ nhiều ảnh nhỏ trong tập ảnh mẫu đó sao cho nhìn một cách khái quát thì vẫn có thể thấy được nội dung của bức ảnh lớn.

3.2 Phân tích, thiết kế

Hoạt động của chương trình :

Bước 1: Đưa vào chương trình một ảnh lớn định dạng jpg, và một tập các ảnh nhỏ có kích thước bằng nhau.

Bước 2: Chương trình sẽ trích chọn đặc trưng về màu sắc của các ảnh nhỏ này

Bước 3: Ảnh nguồn sẽ được chỉnh lại kích thước sao cho kích thước của nó là bội số của kích thước các ảnh nhỏ.

Bước 4: Chương trình chia ảnh lớn thành các ảnh nhỏ có kích thước bằng các ảnh nhỏ được dùng làm mẫu, đồng thời trích chọn đặc trưng của các phần nhỏ sau khi được chia ra từ ảnh lớn rồi đem so sánh với đặc trưng về màu của từng ảnh trong tập ảnh mẫu (sử dụng độ đo Euclid) để chọn ra ảnh tương đồng về màu sắc và thay thế cho nó.

Bước 5: Ghép lại các ảnh nhỏ sau khi được thay thế để tạo ra ảnh khảm.

Chương trình cần phải có một số các chức năng chính như sau:

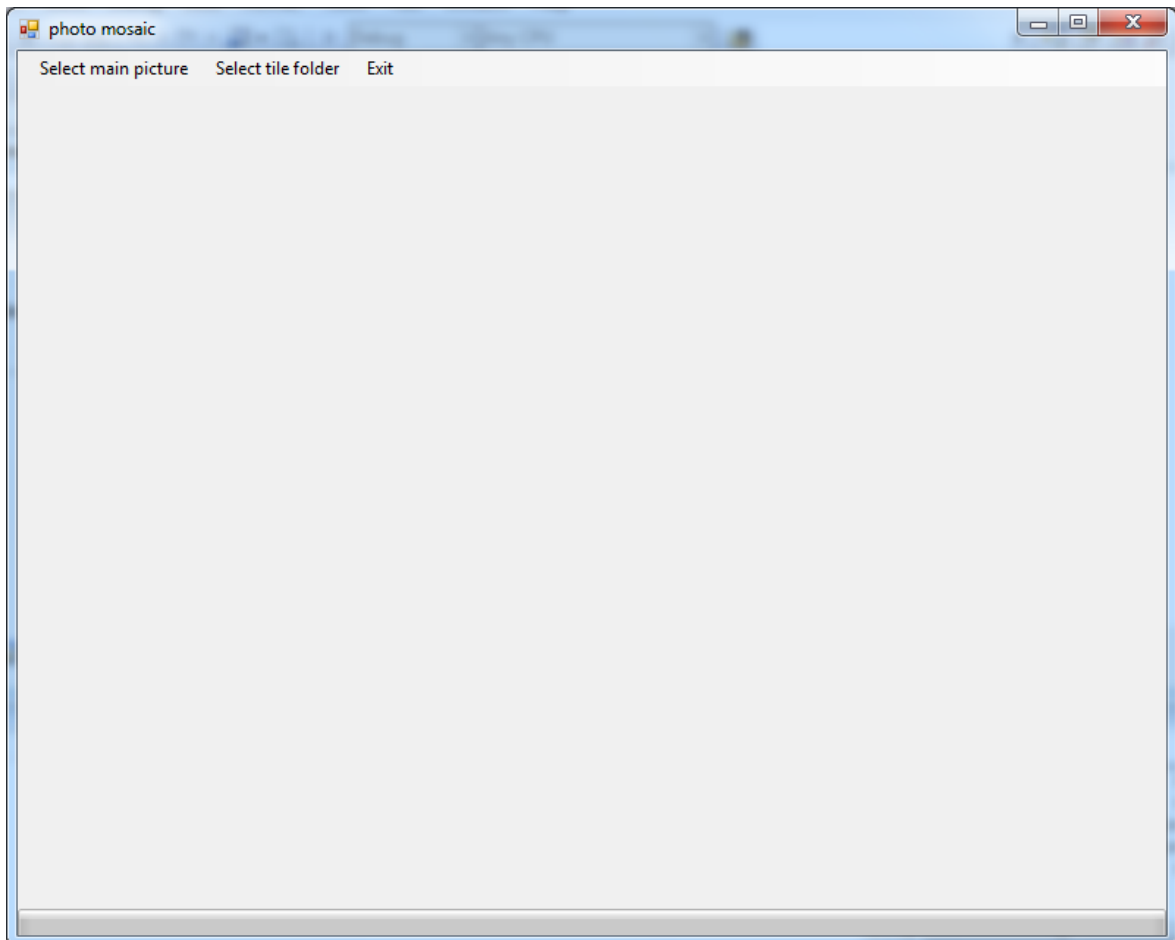
1. Đọc và hiển thị ảnh.
2. Lấy đường dẫn của tập ảnh mẫu.
3. Xử lý và hiển thị kết quả.

3.3 Chương trình khảm ảnh

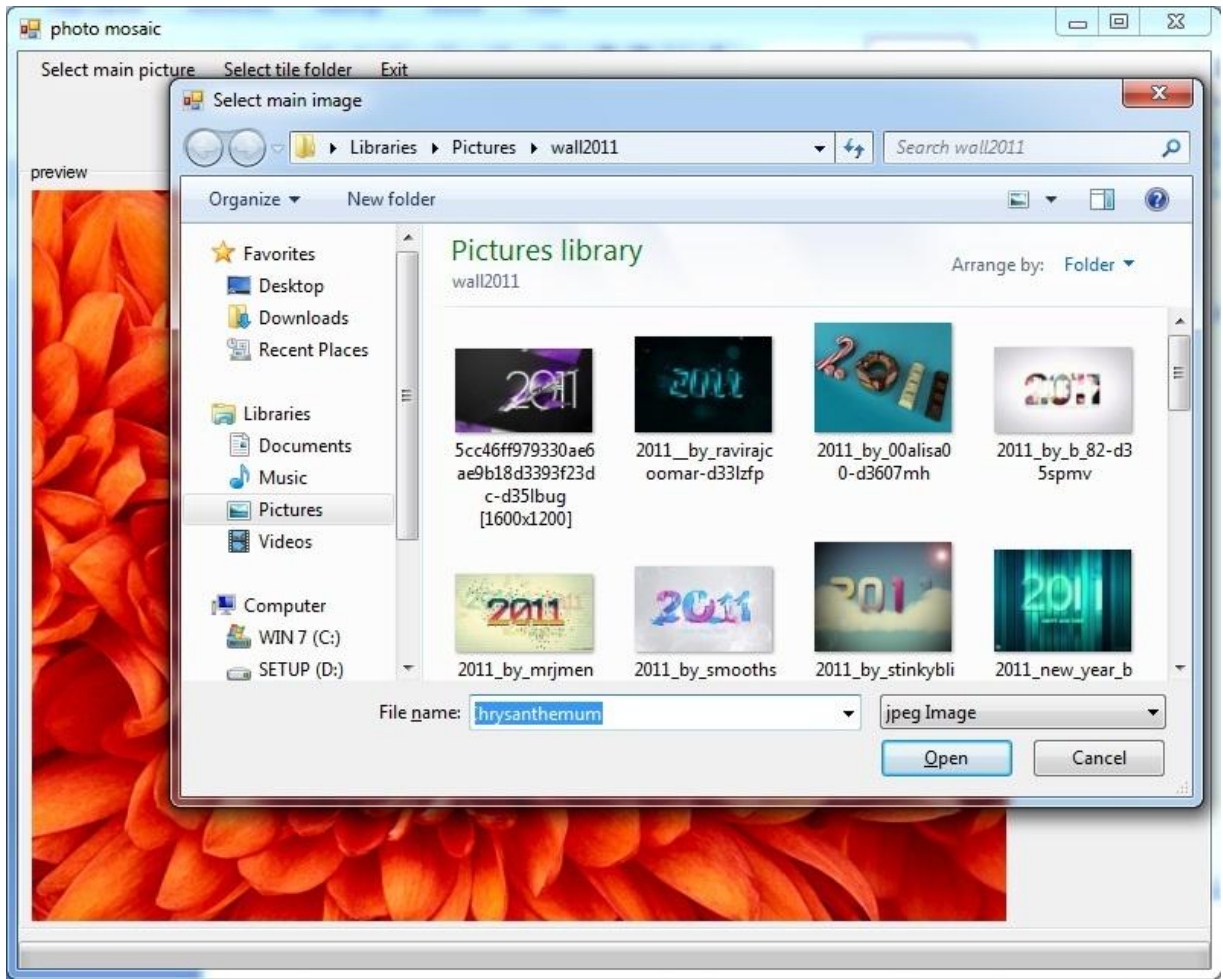
Chương trình được xây dựng bằng bộ công cụ Visual Studio 2008 sử dụng ngôn ngữ lập trình C#. Có 2 tập ảnh mẫu dùng để khảm ảnh gồm 1000 ảnh jpg kích thước 50x50 và 1000 ảnh jpg kích thước 20x20. Chương trình sử dụng phương pháp trích chọn đặc trưng về màu sắc để tìm ra ảnh tương đồng cho các vùng trên ảnh gốc. Nhóm các modul chính của chương trình:

1. Chọn và hiển thị ảnh nguồn dùng để khảm
2. Chọn thư mục chứa ảnh mẫu
3. Xử lý để cho ra ảnh khảm và hiển thị ảnh khảm sau khi được xử lý.

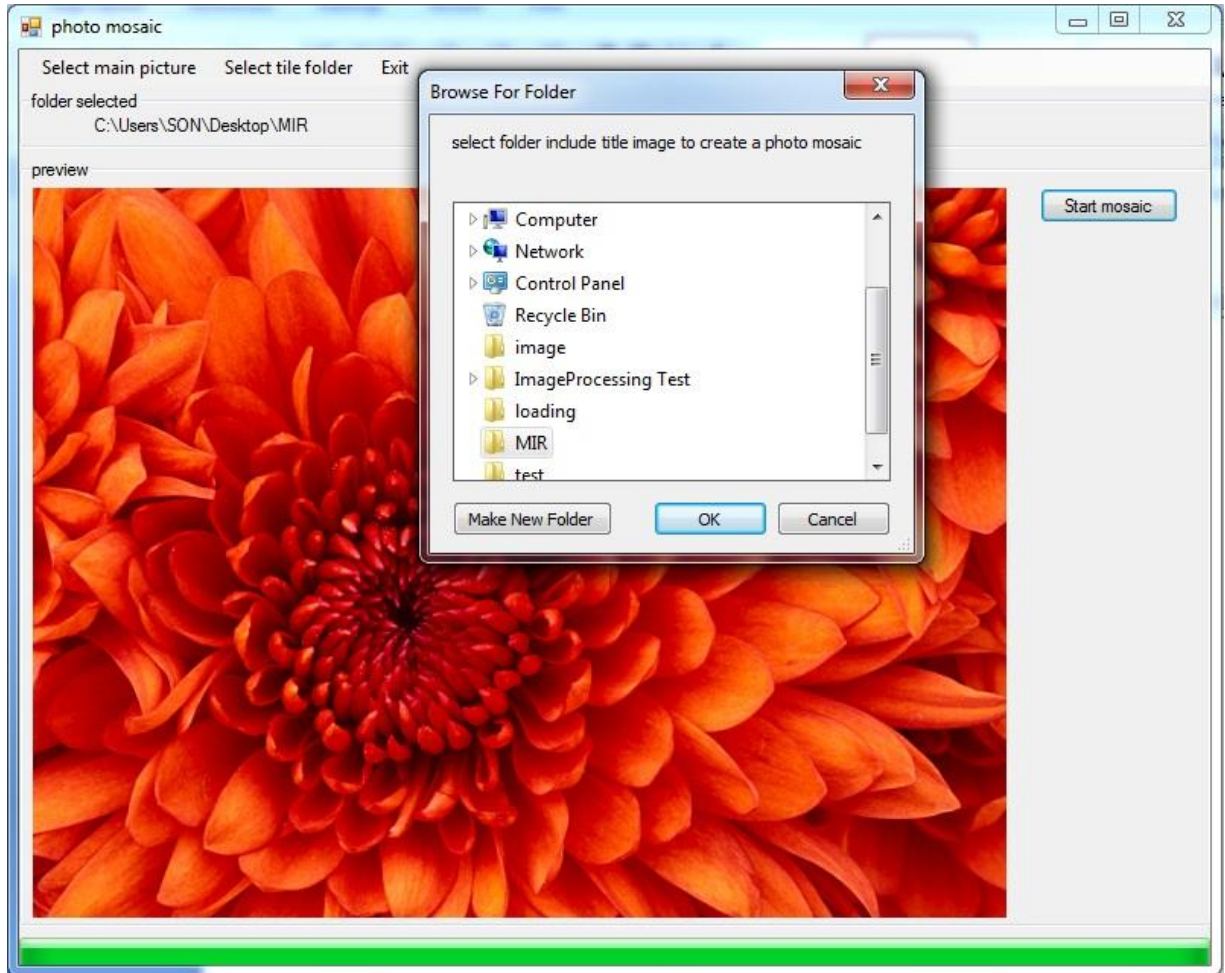
Một số giao diện chính của chương trình:



Hình 3.1: Giao diện chương trình khi khởi động



Hình 3.2: Giao diện chương trình sau khi chọn ảnh nguồn dùng để khám



Hình 3.3: Chọn thư mục chứa tập ảnh mẫu.

Một số kết quả thu được:



Hình 3.4: Ảnh khảm sử dụng ảnh mẫu kích thước 20x20



Hình 3.5: Ảnh khảm sử dụng ảnh mẫu kích thước 50x50



Hình 3.6: Hình ảnh phong cảnh được khám từ tập ảnh mẫu 50x50



Hình 3.7: Hình ảnh phong cảnh được khám từ tập ảnh mẫu 20x20

Qua hình 3.4, 3.5, 3.6, 3.7 có thể thấy chỉ có rất ít những ảnh nhỏ có màu sắc tương đồng với vị trí tương ứng của nó trên ảnh gốc. Điều này chứng tỏ thuật toán và phương pháp tạo hình khảm còn rất nhiều sai sót.

KẾT LUẬN

Ngày nay, cùng với sự phát triển mạnh mẽ về mọi mặt của công nghệ thông tin, truyền thông thì xử lý ảnh cùng với các phạm trù của nó cũng có sự phát triển nhanh chóng và đóng một vai trò quan trọng trong nhiều lĩnh vực. Khảm ảnh chính là một trong các phạm trù đó.

Nội dung đề án đã chỉ ra được một số các khái niệm và kĩ thuật khảm ảnh. Chương trình thử nghiệm khảm ảnh do em xây dựng từ một trong các kĩ thuật tìm hiểu được đã có thể tạo ra ảnh khảm từ nhiều ảnh nhỏ nhưng chương trình chưa được hoàn chỉnh do kết quả thu được không chính xác. Với kết quả thu được như vậy có thể do rất nhiều các nguyên nhân khác nhau và chương trình cần được xem xét lại một cách kĩ lưỡng để có thể hoàn chỉnh hơn.

Mặc dù đã hoàn thành được mục tiêu chính của đề án nhưng do điều kiện về thời gian có hạn mà lĩnh vực cần tìm hiểu cũng tương đối rộng nên những gì tìm hiểu được trong đề án sẽ khó tránh khỏi những thiếu sót. Chương trình thử nghiệm cũng chưa thực sự hoàn thiện nhưng đó cũng là một kết quả khả quan. Trong thời gian tới nếu có điều kiện em sẽ tìm hiểu thêm và có thể sẽ xây dựng một chương trình thử nghiệm về khảm ảnh hoàn chỉnh hơn.

TÀI LIỆU THAM KHẢO

Tài liệu Tiếng Việt

- [1]. Đỗ Năng Toàn, Phạm Việt Bình (2007), *Giáo trình xử lý ảnh*, Nhà xuất bản Đại học Thái Nguyên.
- [2]. Lương Mạnh Bá, Nguyễn Thanh Thủy (2007), *Nhập môn xử lý ảnh số*, Nhà xuất bản KHKT.

Tài liệu Tiếng Anh

- [3]. Lowe David(2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 2004
- [4]. Tee Cheng Siew(2008). Feature selection for content-based image retrieval using statistical discriminant analysis. PhD thesis Faculty of Computer Science and Information System Universiti Teknologi Malaysia. 2008
- [5]. PETER J. BURT and EDWARD H. ADELSONRCA, A Multiresolution Spline With Application to Image Mosaics, David Sarnoff Research Center.

Website

- [6]. <http://msdn.microsoft.com>
- [7]. <http://www.picturemosaics.com>
- [8]. <http://scien.stanford.edu>

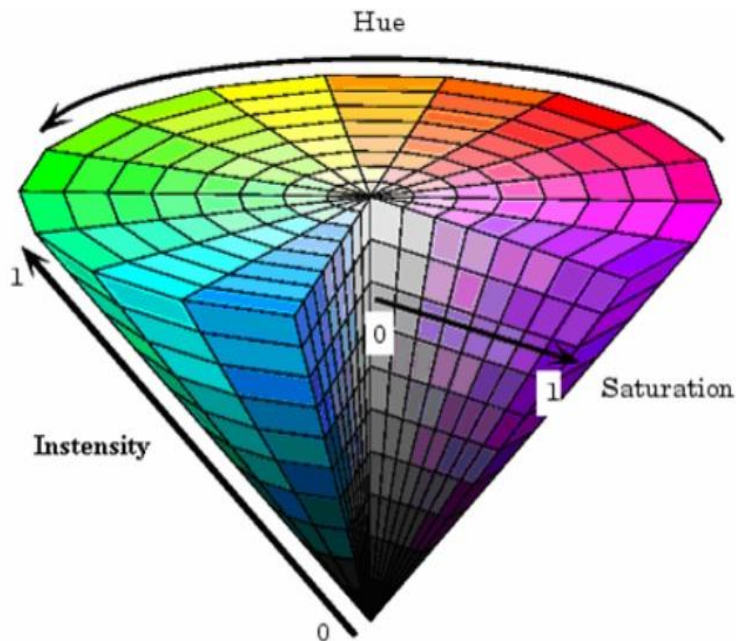
PHỤ LỤC

1. Hệ thống màu HSI

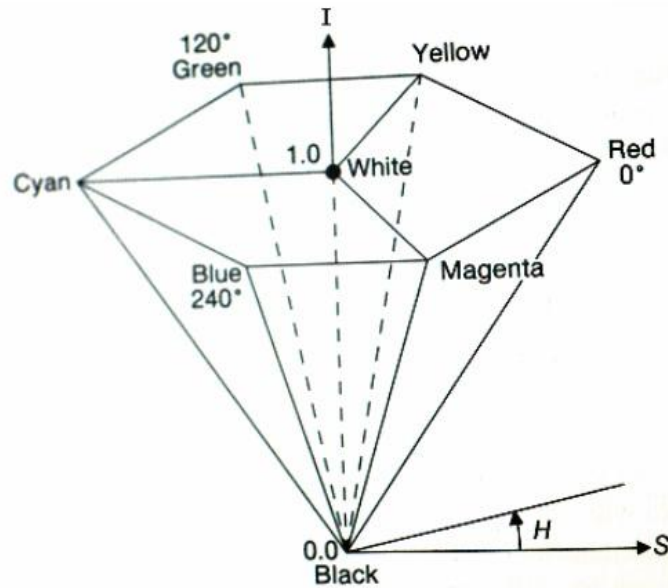
Hệ thống màu HSI mã hóa thông tin màu sắc bằng cách chia giá trị intensity I từ hai giá trị được mã hóa thuộc về độ hội tụ của màu hue H và saturation S. Thành phần không gian màu HSI gồm có ba phần:

- Hue được định nghĩa có giá trị $0-2\pi$, mang thông tin về màu sắc.
- Saturation có giá trị $0-1$, mang giá trị về độ thuần khiết của thành phần Hue.
- Intensity(Value) mang thông tin về độ sáng của điểm ảnh.

Ta có thể hình dung không gian màu HSI như là vật hình nón. Với trục chính biểu thị cường độ sáng Intensity. Khoảng cách đến trục biểu thị độ tập chung Saturation. Góc xung quanh trục biểu thị cho sắc màu Hue. Đôi khi, hệ thống màu HSI được coi như là hệ thống màu HSV dùng Value thay vì Intensity. Hệ thống màu HSI thì thích hợp hơn với một số thiết kế đồ họa bởi vì nó cung cấp sự điều khiển trực tiếp đến ánh sáng và hue. Hệ thống màu HSI cũng hỗ trợ tốt hơn cho những thuật toán xử lý ảnh vì sự tiêu chuẩn hóa về ánh sáng và tập chung vào hai tham số về độ hội tụ màu, và cường độ màu



Hình 1: Hệ màu HIS



Hình 2 : Không gian màu HSI

Hệ thống màu HSI có sự phân chia rõ rệt giữa ánh sáng và màu sắc. Do đó có khả năng rất lớn được áp dụng cho việc tính đặc trưng và so sánh sự giống nhau về màu sắc của hai ảnh. Do đó nó rất thích hợp cho việc tìm kiếm ảnh dựa vào màu.

Sự giống và khác nhau giữa hai ảnh về mặt màu sắc đối với mắt người chỉ mang ý nghĩa tương đối. Do đó khi áp dụng vào bài toán này trên máy tính thì ta cũng giả lập sự tương đối này. Phương pháp chính của việc tìm kiếm theo màu sắc là dùng lược đồ màu để làm đặc trưng cho từng ảnh. Do những đặc điểm riêng của mô hình màu HSI và đặc trưng của việc tìm kiếm nên tính lược đồ màu cũng được dùng một mô hình rất đặc biệt để phù hợp cho những đặc điểm

2. Hệ thống màu CMYK

Từ CMYK (hay đôi khi là YMCK) là từ viết tắt trong tiếng Anh để chỉ mô hình màu loại trừ sử dụng trong in ấn màu. Mô hình màu này dựa trên cơ sở trộn các chất màu của các màu sau:

- C=Cyan trong tiếng Anh có nghĩa là màu xanh lơ
- M=Magenta trong tiếng Anh có nghĩa là màu hồng sẫm
- Y=Yellow trong tiếng Anh có nghĩa là màu vàng

- K=Key (trong tiếng Anh nên hiểu theo nghĩa là cái gì đó *then chốt* hay *chủ yếu* để ám chỉ màu đen mặc dù màu này có tên tiếng Anh là *black* do chữ B đã được sử dụng để chỉ màu xanh lam (blue) trong mô hình màu RGB để tạo các màu khác.

Hỗn hợp của các màu CMY lý tưởng là loại trừ (các màu này khi in cùng một chỗ trên nền trắng sẽ tạo ra màu đen). Nguyên lý làm việc của CMYK là trên cơ sở hấp thụ ánh sáng. Màu mà người ta nhìn thấy là từ phần của ánh sáng không bị hấp thụ. Trong CMYK hồng sẫm cộng với vàng sẽ cho màu đỏ, hồng sẫm cộng với xanh lơ cho màu xanh lam, xanh lơ cộng với vàng sinh ra màu xanh lá cây và tổ hợp của các màu xanh lơ, hồng sẫm và vàng tạo ra màu đen.

Hệ thống màu CMY dường như là một sự đảo ngược của hệ thống màu RGB. Đặc tính của nó là sự đơn giản, ứng dụng nhiều trong thực tế. Tuy nhiên khuyết điểm của nó cũng tương tự như không gian màu RGB, tức là cách mã hóa khác với cách mà con người cảm nhận về màu sắc không tương thích với việc so sánh nội dung các ảnh.